

BUNDESREPUBLIK DEUTSCHLAND



REC'D 13 OCT 2004

WIPO PCT

**Prioritätsbescheinigung über die Einreichung
einer Patentanmeldung**

Aktenzeichen: 10 2004 022 065.4

Anmeldetag: 5. Mai 2004

Anmelder/Inhaber: Degussa AG, 40474 Düsseldorf/DE

Bezeichnung: Screeningverfahren für Hydantoinrazematen

Priorität: 6. Juni 2003 DE 103 26 109.5

IPC: C 12 N, C 12 Q, C 07 H

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

BEST AVAILABLE COPY

Die angehefteten Stücke sind eine richtige und genaue Wiedergabe der ursprünglichen Unterlagen dieser Patentanmeldung.

München, den 27. September 2004
Deutsches Patent- und Markenamt
Der Präsident
Im Auftrag

Dzierzon

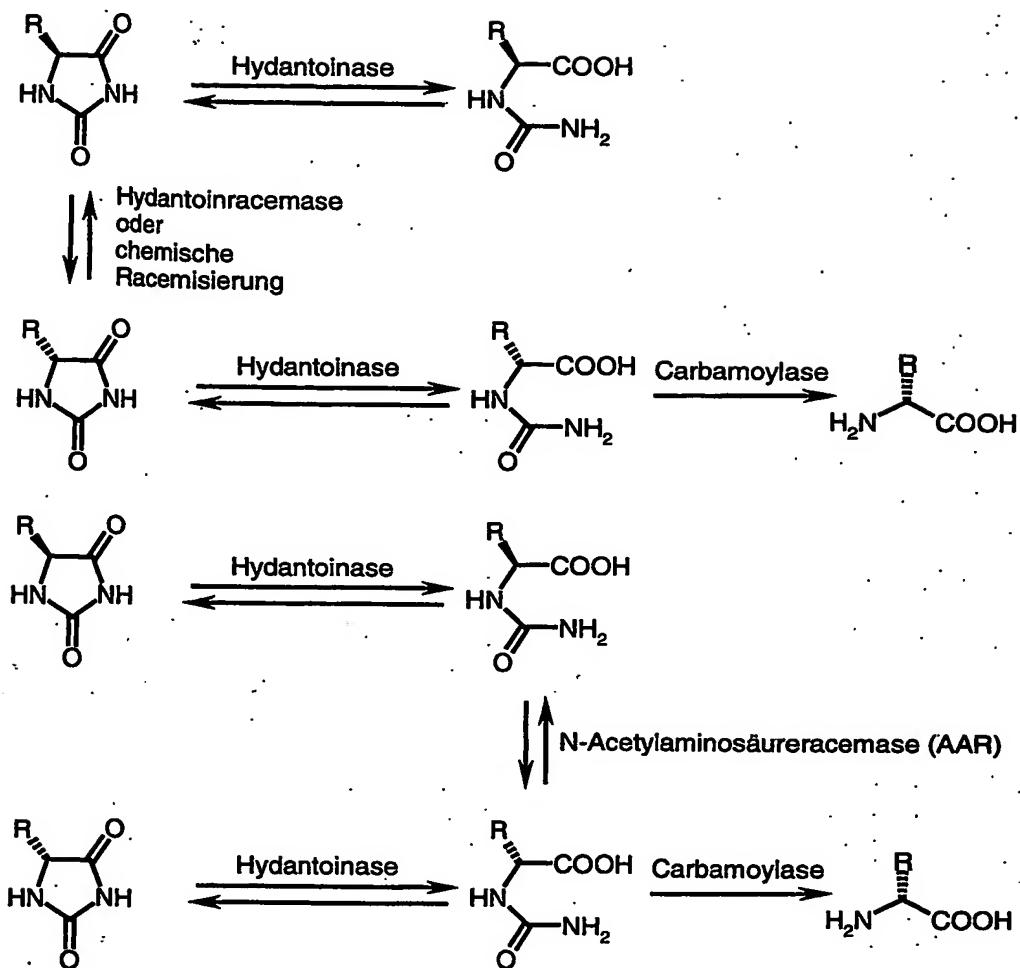
Screeningverfahren für Hydantoinrazematen

Die vorliegende Erfindung ist auf ein Screeningverfahren zur Detektion verbesserter Hydantoinrazematen, neue Hydantoinrazematen selbst und deren Verwendung zur 5 Herstellung von N-Carbamoyl-Aminosäuren gerichtet.

Diese optisch aktiven Verbindungen sind in der organischen Synthese zur Herstellung von z.B. bioaktiven Wirkstoffen häufig eingesetzte Verbindungen. Sie kommen auch in chiralen Auxiliaren z.B. in Form der Aminoalkohole (Evans-Reagenzien) vor. 10

Die enzymatische Hydrolyse von 5-substituierten Hydantoinen zu N-Carbamoyl-Aminosäuren und deren Weiterreaktion zu den entsprechenden enantiomerenangereicherten Aminosäuren ist eine Standardmethode in der organischen Chemie ("Enzyme 15 Catalysis in Organic Synthesis", Eds.: Drauz, Waldmann, VCH, 1st and 2nd Ed.). Die Enantiodifferenzierung kann dabei entweder auf der Stufe der Hydantoinhydrolyse durch Hydantoinasen erfolgen oder aber wahlweise bei der Spaltung der N-Carbamoyl-Aminosäuren mittels enantioselektiver 20 Carbamoylasen. Da die Enzyme nur jeweils eine optische Antipode der entsprechenden Verbindung umsetzen, wird versucht, die andere im Gemisch (in-situ) zu razemisieren, um den vollständigen Umsatz des razemisch leicht herstellbaren Hydantoins in die korrespondierende 25 enantiomerenangereicherte Aminosäure zu gewährleisten. Die Razemisierung kann dabei entweder auf der Stufe der Hydantoina mittels chemischer (Base, Säure, erhöhte Temp.) oder enzymatischer Verfahren erfolgen oder aber auf der Stufe der N-Carbamoyl-Aminosäuren mittels z.B. 30 Acetylaminosäurerazematen (DE10050124) vonstatten gehen. Letztere Variante funktioniert erfolgreich naturgemäß nur bei Einsatz von enantioselektiven Carbamoylasen. Das nachfolgende Schema veranschaulicht diesen Sachverhalt.

Schema 1:



5

Für aromatische Substrate ist die Geschwindigkeit der chemischen Razemisierung der Hydantoine, wie in Tabelle 1 gezeigt, ausreichend hoch, um hohe Raum-Zeit-Ausbeuten für die Herstellung von Aminosäuren nach dem

10 Hydantoinaseverfahren zu gewährleisten. Für aliphatische Hydantoine wie Isobutyl-, Methyl- und Isopropylhydantoin stellt die Razemisierung jedoch einen erheblichen Engpass bei der Synthese aliphatischer Aminosäuren dar.

Tabelle 1: Razemisierungskonstanten von Hydantoinen bei 40°C, pH 8.5 bestimmt durch Anfangsraten gem. einer Reaktion erster Ordnung ($-k_{rac} = \ln([a]/[a_0])$) aus: Hydrolysis and Formation of Hydantoins (Chpt. B 2.4). Syldatk, C. and Pietzsch, M. In: Enzyme catalysis in organic synthesis (Eds.: K. Drauz & H. Waldmann), VCH, 1st and 2nd Ed.).

5'-substituent	k_{rac} (h ⁻¹)	$t_{1/2}$ (h)
Phenyl	2.59	0.27
Methylthioethyl	0.12	5.82
Isobutyl	0.032	21.42
Methyl	0.02	33.98
Isopropyl	0.012	55.90

Dieses Problem zeigt sich beispielsweise bei der in EP759475 beschriebenen Herstellung von enantiomerenangereichertem tert-Butylhydantoin mittels des Hydantoinaseverfahrens. Hier wurden zur vollständigen Umsetzung von 32mM tert.-Butylhydantoin mit 1,5kU R-Hydantoinase 8 Tage bei pH 8,5 und 4 Tage bei pH 9,5 benötigt. Tatsächlich ist die geringe Raum-Zeit-Ausbeute durch die nur langsame chemische Razemisierung von tert-Butylhydantoin ($k_{rac} = 0.009\text{h}^{-1}$ bei 50°C und pH 8.5) bedingt.

Aus dem Stand der Technik sind Hydantoinrazemases aus Mikroorganismen der Gattung *Pseudomonas*, *Mikrobacterium*, *Agrobacterium* und *Arthrobacter* bekannt (Lit.: JP04271784; EP1188826; Cloning and characterization of genes from *Agrobacterium* sp. IP I-671 involved in hydantoin degradation. Hils, M.; Muench, P.; Altenbuchner, J.; Syldatk, C.; Mattes, R. Applied Microbiology and Biotechnology (2001), 57(5-6), 680-688; A new razemase for 5-monosubstituted hydantoins. Pietzsch, Markus;

Syldatk, Christoph; Wagner, Fritz. Ann. N. Y. Acad. Sci. (1992), 672 (Enzyme Engineering XI), 478-83. Lickefett, Holger; Krohn, Karsten; Koenig, Wilfried A.; Gehrcke, Barbel; Syldatk, Christoph. Tetrahedron: Asymmetry (1993),

5 4(6), 1129-35; Purification and characterization of the hydantoin razemase of *Pseudomonas* sp. strain NS671 expressed in *Escherichia coli*. Watabe, Ken; Ishikawa, Takahiro; Mukohara, Yukuo; Nakamura, Hiroaki. J. Bacteriol. (1992), 174(24), 7989-95).

10 Von den Hydantoinrazemasesen aus *Arthrobacter aurescens* DSM 3745, *Pseudomonas* sp. NS671 und *Microbacterium liquefaciens* ist bekannt, dass diese Enzyme aliphatische Hydantoine wie beispielsweise Isopropylhydantoin oder Isobutylhydantoin nur schwach razemisieren. Darüber hinaus weiß man, dass die

15 Hydantoinrazemasesen aus *Arthrobacter aurescens* DSM 3747 aromatische Hydantoine wie Indolylmethylhydantoin oder Benzylhydantoin bevorzugt, wohingegen aliphatische Hydantoine wie Methylthioethylhydantoin vergleichsweise schwach oder im Fall von Isopropylhydantoin überhaupt nicht umgesetzt werden (A new razemase for 5-monosubstituted hydantoins. Pietzsch, Markus; Syldatk, Christoph; Wagner, Fritz. Ann. N. Y. Acad. Sci. (1992), 672 (Enzyme Engineering XI), 478-83.).

Die niedrige Aktivität von Hydantoinrazemasesen begrenzt daher häufig das wirtschaftliche Potential dieser Route.

Um in geeigneter Zeit möglichst viele Hydantoinrazemasesen auf ihr Potential zur Razemisierung von aliphatischen Hydantoine prüfen zu können, lag die Aufgabe der vorliegenden Erfindung unter anderem in der Angabe eines geeigneten Screeningverfahrens für Hydantoinrazemasesen.

30 Darüber hinaus sollte das erfindungsgemäße Screeningverfahren als Bestandteil für ein Mutagenseverfahren zur Gewinnung neuer und besserer Hydantoinrazemasesen einsetzbar sein. Ebenfalls Aufgabe der vorliegenden Erfindung war die Angabe neuer

Hydantoinrazemosen, die den Hydantoinrazemosen des Standes der Technik zumindest in Selektivität und/oder Aktivität und/oder Stabilität überlegen sind.

Diese Aufgabe wird anspruchsgemäß gelöst. Anspruch 1
5 bezieht sich auf ein Screeningverfahren für Hydantoinrazemosen. Unteransprüche 2 bis 4 zeigen vorteilhafte Ausführungsformen des Screeningverfahrens auf. Anspruch 5 beschäftigt sich mit einem Mutageneseverfahren zur Herstellung neuer Hydantoinrazemosen unter Anwendung 10 des erfindungsgemäßen Screeningverfahrens. Ansprüche 6 bis 11 beziehen sich auf neue Hydantoinrazemosen sowie die sie codierenden Nukleinsäuresequenzen und deren Verwendung. Ansprüche 12 bis 14 richten sich auf Vehikel, welche die 15 erfindungsgemäßen Hydantoinrazemosen aufweisen, bzw. spezielle Primer für deren Herstellung.

Dadurch, dass man ein Screeningverfahren für Hydantoinrazemosen angibt, bei dem man
a) eine enantioselektive Hydantoinase und
b) die zu prüfende Hydantoinrazemase, welche eine 20 verglichen mit der Hydantoinase unter a) langsamere Umsetzungsrate aufweist, auf
c) ein chirales Hydantoin einwirken lässt, welches in zur Selektivität der Hydantoinase entgegengesetzter enantiomerenangereicherter Form eingesetzt wird, und
d) die resultierende N-Carbamoyl-Aminosäure oder die 25 freigesetzten Protonen zeitabhängig detektiert, gelangt man überraschend einfach und dennoch vorteilhaft zu einer Möglichkeit, viele Hydantoinrazemosen in kurzer Zeit auf ihre Fähigkeit hin zu überprüfen, in verbesserter Weise 30 Hydantoinrazemisieren zu können.

Durch Einsatz eines L-Enantiomers eines 5'-monosubstituierten Hydantoins und Verwendung einer D-selektiven Hydantoinase, welche aufgrund ihrer 35 Enantioselektivität bevorzugt dass entstehende D-Enantiomer

des Hydantoins schnell hydrolysiert, kann durch die Bildung der N-Carbamoyl-D-Aminosäure oder freiwerdende Protonen die Razemisierungsgeschwindigkeit und damit die Aktivität der Hydantoinrazemase auf einfache Weise gemessen werden. Die

5 Quantifizierung der N-Carbamoyl-Aminosäure kann dabei durch dem Fachmann bekannte Methoden wie beispielsweise HPLC oder colorimetrische Methoden erfolgen. Die Quantifizierung über Protonen kann auf einfache Weise über pH Indikatoren, bevorzugt Cresol Rot, erfolgen. Es sei darauf hingewiesen,

10 dass in dem Verfahren sowohl D- als auch L-Enantiomere von Hydantoinen mit unterschiedlichen ggf. aliphatischen 5'-Substituenten eingesetzt werden können. Beim Einsatz der D-Hydantoine sind dementsprechenden L-selektive Hydantoinasen im Screeningverfahren einzusetzen.

15 Im erfindungsgemäßen Verfahren eingesetzt werden vorteilhaft aliphatische in 5'-Stellung substituierte Hydantoine. Unter aliphatisch substituierten Hydantoinen wird in diesem Zusammenhang ein System verstanden, welches in 5'-Stellung an dem Hydantoinheterozyklus einen Rest

20 aufweist, der über ein C-Atom mit sp^3 -Hybridisierung an den Heterozyklus gebunden ist. Bevorzugte 5'-Substituenten sind dabei Methyl, Ethyl, Butyl, Propyl, tertiar-Butyl, Isopropyl und Isobutyl. Ganz besonders bevorzugt ist Ethyl-Hydantoin.

25 Als Hydantoinasen können sämtliche in der Literatur bekannten Hydantoinasen eingesetzt werden, welche das über die Hydantoinrazemase gebildete Enantiomer des Hydantoins enantioselektiv hydrolysiieren, wobei diese Hydrolyse schneller als die Razemisierungsgeschwindigkeit sein muss.

30 Bevorzugte Hydantoinasen sind dabei die kommerziellen Hydantoinasen 1 & 2 von Roche, die Hydantoinasen der Gattungen *Agrobacterium*, *Arthrobacter*, *Bacillus*, *Pseudomonas*, *Flavobacterium*, *Pasteurella*, *Microbacterium*, *Vigna*, *Ochrobactrum*, *Methanococcus*, *Burkholderia* und

35 *Streptomyces*. (Hils, M.; Muench, P.; Altenbuchner, J.;

Syldatk, C.; Mattes, R. Cloning and characterization of genes from *Agrobacterium* sp. IP I-671 involved in hydantoin degradation. *Applied Microbiology and Biotechnology* (2001), 57(5-6), 680-688.

5 Soong, C.-L.; Ogawa, J.; Shimizu, S. Cyclic ureide and imide metabolism in microorganisms producing a D-hydantoinase useful for D-amino acid production. *Journal of Molecular Catalysis B: Enzymatic* (2001), 12(1-6), 61-70.

10 Wiese, Anja; Wilms, Burkhard; Syldatk, Christoph; Mattes, Ralf; Altenbuchner, Josef. Cloning, nucleotide sequence and expression of a hydantoinase and carbamoylase gene from *Arthrobacter aurescens* DSM 3745 in *Escherichia coli* and comparison with the corresponding genes from *Arthrobacter aurescens* DSM 3747. *Applied Microbiology and Biotechnology* (2001), 55(6), 750-757.

15 Yin, Bang-Ding; Chen, Yi-Chuan; Lin, Sung-Chyr; Hsu, Wen-Hwei. Production of D-amino acid precursors with permeabilized recombinant *Escherichia coli* with D-hydantoinase activity. *Process Biochemistry* (Oxford) (2000), 35(9), 915-921.

20 Park, Joo-Ho; Kim, Geun-Joong; Lee, Seung-Goo; Lee, Dong-Cheol; Kim, Hak-Sung. Purification and characterization of thermostable D-hydantoinase from *Bacillus thermocatenulatus* GH-2. *Applied Biochemistry and Biotechnology* (1999), 81(1), 53-65; Pozo, C.; Rodelas, B.; de la Escalera, S.; Gonzalez-Lopez, J. D,L-Hydantoinase activity of an *Ochrobactrum anthropi* strain. *Journal of Applied Microbiology* (2002), 92(6), 1028-1034; Chung, Ji Hyung; Back, Jung Ho; Lim, Jae-Hwan; Park, Young In; Han, Ye Sun. Thermostable hydantoinase from a hyperthermophilic archaeon, *Methanococcus jannaschii*. *Enzyme and Microbial Technology* (2002), 30(7), 867-874; Xu, Zhen; Jiang, Weihong; Jiao, Ruishen; Yang, Yunliu. Cloning, sequencing and high expression in *Escherichia coli* of D-hydantoinase gene from *Burkholderia pickettii*. *Shengwu Gongcheng Xuebao* (2002), 18(2), 149-154; Las Heras-Vazquez, Francisco Javier; Martinez-Rodriguez, Sergio; Mingorance-Cazorla, Lydia; Clemente-Jimenez, Josefa Maria; Rodriguez-Vico, Felipe.

Overexpression and characterization of hydantoin racemase from *Agrobacterium tumefaciens* C58. Biochemical and Biophysical Research Communications (2003), 303(2), 541-547; DE 3535987; EP 1275723; US 6087136; WO 0281626; US 5 2002045238; DE 4328829; WO 9400577; WO 9321336; JP 04325093; NL 9001680; JP 2003024074; WO 0272841; WO 0119982; WO 9620275).

10 Ganz besonders bevorzugt ist die Verwendung der Hydantoinase aus *Arthrobacter crystallopoietes*, insbesondere der aus DSM 20117.

Wie schon angedeutet sollte die Umsetzungsgeschwindigkeit der Hydantoinase die der Razemase übertreffen. Vorzugsweise liegt das Verhältnis der Geschwindigkeitskonstanten der Hydantoinase zur Hydantoinrazemase (k_{Hyd}/k_{Raz}) bei >2, 15 besonders bevorzugt bei > 10 und ganz besonders bevorzugt bei >50.

Gegenstand der Erfindung ist ebenfalls ein Verfahren zur Herstellung von verbesserten Hydantoinrazemassen, welches sich dadurch auszeichnet, dass man 20 a) die Nukleinsäuresequenz codierend für die Hydantoinrazemase einer Mutagenese unterwirft, b) die aus a) erhältlichen Nukleinsäuresequenzen in einen geeigneten Vektor kloniert und diesen in ein geeignetes Expressionssystem transferiert und 25 c) die gebildeten Hydantoinrazemassen mit verbesserter Aktivität und/oder Selektivität und/oder Stabilität mittels eines erfindungsgemäßen Screeningverfahrens detektiert und isoliert.

Als Ausgangsgene für die Mutagenese der Hydantoinrazemassen 30 können sämtliche bekannten und in der angeführten Literatur erwähnten Hydantoinrazemasegene dienen. Bevorzugt sind dabei die Hydantoinrazemasegene von *Arthrobacter*, *Pseudomonas*, *Agrobacterium* und *Micrococcus* (Wiese A; Pietzsch M; Syldatk C; Mattes R; Altenbuchner J Hydantoin 35 racemase from *Arthrobacter aurescens* DSM 3747: heterologous

expression, purification and characterization. JOURNAL OF BIOTECHNOLOGY (2000 Jul 14), 80(3), 217-30; Watabe K; Ishikawa T; Mukohara Y; Nakamura H Purification and characterization of the hydantoin racemase of *Pseudomonas* sp. strain NS671 expressed in *Escherichia coli*. JOURNAL OF BACTERIOLOGY (1992 Dec), 174(24), 7989-95; Las Heras-Vazquez, Francisco Javier; Martinez-Rodriguez, Sergio; Mingorance-Cazorla, Lydia; Clemente-Jimenez, Josefa Maria; Rodriguez-Vico, Felipe. Overexpression and characterization of hydantoin racemase from *Agrobacterium tumefaciens* C58. Biochemical and Biophysical Research Communications (2003), 303(2), 541-547; EP. 1188826). Ganz besonders bevorzugt ist das Hydantoinrazemasegen aus *Arthrobacter aurescens* welches für die Proteinsequenz in Seq.ID.Nr. 2 codiert.

Zur Mutagenese der Hydantoinrazemase können sämtliche in der Literatur bekannten Methoden wie beispielsweise Zufallsmutagenese, Sättigungsmutagenes, Kassetten-Mutagenese oder Rekombinationsmethoden verwendet werden (May, Oliver; Voigt, Christopher A.; Arnold, Frances H. Enzyme engineering by directed evolution. Enzyme Catalysis in Organic Synthesis (2nd Edition) (2002), 1 95-138; Bio/Technology 1991, 9, 1073-1077; Horwitz, M. und Loeb, L., Promoters Selected From Random DNA-Sequences, Proc Natl Acad Sci USA 83, 1986, 7405-7409; Dube, D. und L. Loeb, Mutants Generated By The Insertion Of Random Oligonucleotides Into The Active-Site Of The Beta-Lactamase Gene, Biochemistry 1989, 28, 5703-5707; Stemmer, P.C., Rapid evolution of a protein in vitro by DNA shuffling, Nature 1994, 370, 389-391 und Stemmer, P.C., DNA shuffling by random fragmentation and reassembly: In vitro recombination for molecular evolution. Proc Natl Acad Sci USA 91, 1994, 10747-10751).

Die Klonierung und Expression kann wie in der weiter unten angegebenen Literatur durchgeführt werden. Das Verfahren kann mehrmals hintereinander ggf. mit wechselnden Mutagenesestrategien durchgeführt werden.

Gegenstand der Erfindung sind ebenfalls rec-Polypeptide oder die diese codierende Nukleinsäuresequenzen, welche nach dem eben genannten Mutageneseverfahren erhältlich sind.

5 Ebenso ein Aspekt der Erfindung ist die Verwendung der so hergestellten Polypeptide zur Herstellung von chiralen enantiomerenangereicherten N-Carbamoyl-Aminosäuren oder Aminosäuren. Die erfindungsgemäß hergestellten Nukleinsäuresequenzen können zur Herstellung von
10 Ganzzellkatalysatoren dienen.

Einen Teil der vorliegenden Erfindung bilden auch Hydantoinrazematen, welche in Position 79 einen Aminosäureaustausch mit einer Aminosäure ausgewählt aus der Gruppe bestehend aus A, R, N, D, C, Q, E, H, I, L, K, M, F, P, S, T, Y oder V aufweisen. Interessant ist, dass die Aminosäuren, welche diese Position umgeben, für viele Hydantoinrazematen vollständig konserviert sind. Die Konsensussequenz lautet: FX₁DX₂GL (Seq.ID.Nr. 1), wobei X₂ P oder T darstellt und X₁ W oder G darstellt. Bevorzugte
15 Mutanten weisen daher die oben genannte Konsensussequenz auf, wobei X₁ vorzugsweise eine Aminosäure ausgewählt aus der Gruppe bestehend aus A, R, N, D, C, Q, E, H, I, L, K, M, F, P, S, T, Y oder V darstellt. X₁ entspricht dabei der Position 79. Bevorzugte Mutanten sind in Tabelle 2
20 dargestellt.
25

Tabelle 2:

Mutanten Name	Mutation (codon)	Mutation X ₁ (Aminosäure)	Aktivitätsänderung	Seq. ID Nr.
3CH11	GGG → GAG	G79E	2	5
1BG7	GGG → AGG	G79R	2	3
BB5	GGG → TTG	G79L	4	9
AE3	GGG → CAG	G79Q	4	7

Weitere äußerst vorteilhafte Kombinationen von X₁ und X₂

5 Hydantoinrazemase sind in folgender Tabelle 3 aufgeführt.

Tabelle 3: Vorteilhafte Kombinationen von X₁ und X₂ in dem Konsensusmotiv FX₁DX₂GL

X ₁	L	E	Q	R	L	E	Q	R
X ₂	P	P	P	P	T	T	T	T

10 Von besonderem Vorteil ist es, wenn die Hydantoinrazemase die oben angegebene Konsensusregion und zusätzlich eine Homologie von >40% zur Hydantoinrazemase aus DSM 20117 aufweisen.

Weiterhin Gegenstand der Erfindung sind isolierte Nukleinsäuresequenzen codierend für eine Hydantoinrazemase 15 ausgewählt aus der Gruppe:

- einer Nukleinsäuresequenz codierend für eine erfindungsgemäße Hydantoinrazemase,
- einer Nukleinsäuresequenz, die unter stringenten Bedingungen mit der Nukleinsäuresequenz codierend für

eine erfindungsgemäße Hydantoinrazemase oder der dazu komplementären Sequenz hybridisiert,

- c) einer Nukleinsäuresequenz gemäß den Seq.ID.Nr. 3, 5, 7 oder 9 oder solchen mit einer Homologie von > 80% zu diesen,
- d) einer Nukleinsäuresequenz aufweisend 15 aufeinanderfolgende Nukleotide der Sequenzen Seq.ID.Nr. 3, 5, 7 oder 9.

In Bezug auf Punkt d) ist es bevorzugt, wenn die erfindungsgemäße Nukleinsäuresequenz 20, mehr bevorzugt 25, weiter bevorzugt 30, 31, 32, 33, 34 und äußerst bevorzugt mehr als 34 identische konsekutive Nukleinsäuren der Sequenzen Seq.ID.Nr. 3, 5, 7 oder 9 aufweist.

Wie gesagt sind von der Erfindung auch Nukleinsäuresequenzen mitumfasst, welche unter stringenten Bedingungen mit den erfindungsgemäßen einzelsträngigen Nukleinsäuresequenzen oder deren komplementären einzelsträngigen Nukleinsäuresequenzen hybridisieren (b) oder solche, die sich in Sequenzabschnitten gleichen (d).

Als solche sind z.B. spezielle Gensonden oder die für eine PCR notwendigen Primer anzusehen.

Eine Kopplung von Hydantoinrazemase und Hydantoinase und ggf. Carbamoylase kann dabei durch Zusammengeben der freien bzw. immobilisierten Enzyme erfolgen. Bevorzugt ist jedoch, wenn die Hydantoinase gemeinsam mit der Hydantoinrazemase und/oder der Carbamoylase in der selben Zelle exprimiert wird (Ganzzellkatalysator).

Die erfindungsgemäßen Nukleinsäuresequenzen können daher als Bestandteil eines Gens in analoger Weise wie in DE10234764 und dort zitierter Literatur in einen Ganzzellkatalysator kloniert werden.

Sofern dieser dann auch Gene für eine Hydantoinase und/oder Carbamoylase aufweist, ist er im Stande racemische Hydantoine zur Gänze in enantiomerenangereicherte Aminosäuren umzuwandeln. Ohne ein kloniertes

Carbamoylasegen stoppt die Reaktion auf der Stufe der N-Carbamoyl-Aminosäuren.

Vorzugsweise wird ein Organismus wie in der DE10155928 genannt als Wirtsorganismus eingesetzt. Der Vorteil eines

5 derartigen Organismus ist die gleichzeitige Expression aller beteiligten Enzyme, womit nur noch ein rec-Organismus für die Gesamtreaktion angezogen werden muss.

Um die Expression der Enzyme im Hinblick auf ihre Umsetzungsgeschwindigkeiten abzustimmen, können die 10 entsprechenden codierenden Nukleinsäuresequenzen in unterschiedliche Plasmide mit unterschiedlichen Kopienzahlen kloniert und/oder unterschiedlich starke Promotoren für eine unterschiedlich starke Expression der Nukleinsäuresequenzen verwendet werden. Bei derart 15 abgestimmten Enzymsystemen tritt vorteilhafterweise eine Akkumulation einer ggf. inhibierend wirkenden Zwischenverbindung nicht auf und die betrachtete Reaktion kann in einer optimalen Gesamtgeschwindigkeit ablaufen.

Dies ist dem Fachmann jedoch hinlänglich bekannt

20 (Gellissen, G.; Piontek, M.; Dahlems, U.; Jenzelewski, V.; Gavagan, J. W.; DiCosimo, R.; Anton, D. L.; Janowicz, Z. A. (1996), Recombinant *Hansenula polymorpha* as a biocatalyst. Coexpression of the spinach glycolate oxidase (GO) and the *S. cerevisiae* catalase T (CTT1) gene, *Appl. Microbiol.*

Biotechnol.

46, 46-54; Farwick, M.; London, M.; Dohmen, J.; Dahlems, U.; Gellissen, G.; Strasser, A. W.; DE19920712). Die Herstellung eines derartigen Ganzzellkatalysators ist dem Fachmann hinlänglich bekannt (Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), *Molecular cloning: a laboratory manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York; Balbas, P. und Bolivar, F. (1990), *Design and construction of expression plasmid vectors in E.coli, Methods Enzymol.* 185, 14-37; Rodriguez, R.L. und Denhardt, D. T (eds) (1988), *Vectors: a survey of molecular cloning vectors and their uses*, 205-225, Butterworth, Stoneham).

In einer nächsten Ausgestaltung bezieht sich die Erfindung auf Plasmide oder Vektoren aufweisend eine oder mehrere der erfindungsgemäßen Nukleinsäuresequenzen.

Als Plasmide oder Vektoren kommen im Prinzip alle dem Fachmann für diesen Zweck zur Verfügung stehenden Ausführungsformen in Frage. Derartige Plasmide und Vektoren können z. B. von Studier und Mitarbeiter (Studier, W. F.; Rosenberg A. H.; Dunn J. J.; Dubendorff J. W.; (1990), Use of the T7 RNA polymerase to direct expression of cloned

genes, Methods Enzymol. 185, 61-89) oder den Broschüren der Firmen Novagen, Promega, New England Biolabs, Clontech oder Gibco BRL entnommen werden. Weiter bevorzugte Plasmide und Vektoren können gefunden werden in: Glover, D. M. (1985), DNA cloning: a practical approach, Vol. I-III, IRL Press Ltd., Oxford; Rodriguez, R.L. und Denhardt, D. T (eds) (1988), Vectors: a survey of molecular cloning vectors and their uses, 179-204, Butterworth, Stoneham; Goeddel, D. V. (1990), Systems for heterologous gene expression, Methods Enzymol. 185, 3-7; Sambrook, J.; Fritsch, E. F. und Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York.

Plasmide, mit denen das die erfindungsgemäße Nukleinsäure aufweisende Genkonstrukt in ganz bevorzugter Weise in den Wirtsorganismus kloniert werden kann, sind Derivate von pUC18 und pUC19 (Roche Biochemicals), pKK-177-3H (Roche Biochemicals), pBTac2 (Roche Biochemicals), pKK223-3 (Amersham Pharmacia Biotech), pKK-233-3 (Stratagene) oder pET (Novagen). Weitere bevorzugte Plasmide sind pBR322

(DSM3879), pACYC184 (DSM4439) und pSC101 (DSM6202), welche von der DSMZ-Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH, Braunschweig, Germany bezogen werden können.

Als bevorzugt anzusehende Plasmide Gleichfalls ist die Erfindung auf Mikroorganismen aufweisend eine oder mehrere erfindungsgemäße Nukleinsäuresequenzen gerichtet.

Der Mikroorganismus, in den die die erfindungsgemäßen

Nukleinsäuresequenzen enthaltenen Plasmide kloniert werden, dient zur Vermehrung und Gewinnung einer ausreichenden Menge des rekombinanten Enzyms. Die Verfahren hierfür sind dem Fachmann wohlbekannt (Sambrook, J.; Fritsch, E. F. und

5 Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York). Als Mikroorganismen können im Prinzip alle dem Fachmann für diesen Zweck in Frage kommenden Organismen wie z.B. Hefen wie *Hansenula polymorpha*, *Pichia* sp.,

10 *Saccharomyces cerevisiae*, Prokaryonten, wie *E. coli*, *Bacillus subtilis* oder Eukaryonten, wie Säugerzellen, Insektenzellen herangezogen werden. Vorzugsweise sind *E. coli*-Stämme für diesen Zweck zu benutzen. Ganz besonders bevorzugt sind: *E. coli* XL1 Blue, W3110, DSM14459

15 (PCT/US00/08159), NM 522, JM101, JM109, JM105, RR1, DH5 α , TOP 10⁺ oder HB101. Plasmide, mit denen das die erfindungsgemäße Nukleinsäure aufweisende Genkonstrukt vorzugsweise in den Wirtsorganismus kloniert wird, sind weiter oben angegeben.

20 Ein folgender Aspekt der Erfindung richtet sich auf Primer zur Herstellung der erfindungsgemäßen Gensequenzen mittels aller Arten von PCR. Mitumfasst sind die Sense- und Antisense-Primer codierend für die entsprechenden Aminosäuresequenzen, bzw. komplementären DNA-Sequenzen. Geeignete Primer können prinzipiell nach dem Fachmann bekannten Verfahren gewonnen werden. Das Auffinden der erfindungsgemäßen Primer erfolgt durch Vergleich mit bekannten DNA-Sequenzen oder durch Übersetzung der ins Auge gefassten Aminosäuresequenzen in das bevorzugte Codon des

30 betrachteten Organismus (z.B. für *Streptomyces*: Wright F. und Bibb M. J. (1992), Codon usage in the G+C-rich *Streptomyces* genome, Gene 113, 55-65). Gemeinsamkeiten in der Aminosäuresequenz von Proteinen von sogenannten Superfamilien sind hierfür ebenfalls von Nutzen (Firestone, S. M.; Nixon, A. E.; Benkovic, S. J. (1996), Threading your way to protein function, Chem. Biol. 3, 779-783). Weitere

35

Informationen diesbezüglich können gefunden werden in Gait, M. J. (1984), Oligonucleotide synthesis: a practical approach, IRL Press Ltd., Oxford; Innis, M. A.; Gelfound, D. H.; Sninsky, J. J. und White, T.J. (1990), PCR

5 Protocols: A guide to methods and applications, Academic Press Inc., San Diego.

Bevorzugte Primer sind die der Seq.ID.Nr. 11 und 12.

Für die Anwendung können die betrachteten Enzyme (Hydantoinrazemase, Hydantoinasen und/oder Carbamoylasen)

10 wie schon angedeutet in freier Form als homogen aufgereinigte Verbindungen oder als rekombinant (rec-) hergestelltes Enzym verwendet werden. Weiterhin können die Enzyme auch als Bestandteil eines intakten Gastorganismus eingesetzt werden oder in Verbindung mit der 15 aufgeschlossenen und beliebig hoch aufgereinigten Zellmasse des Wirtsorganismus.

Möglich ist ebenfalls die Verwendung der Enzyme in immobilisierter Form (Sharma B. P.; Bailey L. F. und Messing R. A. (1982), Immobilisierte Biomaterialien -

20 Techniken und Anwendungen, Angew. Chem. 94, 836-852). Vorteilhafterweise erfolgt die Immobilisierung durch Lyophilisation (Paradkar, V. M.; Dordick, J. S. (1994), Aqueous-Like Activity of α -Chymotrypsin Dissolved in Nearly Anhydrous Organic Solvents, J. Am. Chem. Soc. 116, 5009-5010; Mori, T.; Okahata, Y. (1997), A variety of lipi-coated glycoside hydrolases as effective glycosyl transfer catalysts in homogeneous organic solvents, Tetrahedron Lett. 38, 1971-1974; Otamiri, M.; Adlercreutz, P.; Matthiasson, B. (1992), Complex formation between

30 chymotrypsin and ethyl cellulose as a means to solbilize the enzyme in active form in toluene, Biocatalysis 6, 291-305). Ganz besonders bevorzugt ist die Lyophilisation in Gegenwart von oberflächenaktiven Substanzen, wie Aerosol OT oder Polyvinylpyrrolidon oder Polyethylenglycol (PEG) oder 35 Brij 52 (Diethylenglycol-mono-cetylether) (Kamiya, N.; Okazaki, S.-Y.; Goto, M. (1997), Surfactant-horseradish

peroxidase complex catalytically active in anhydrous benzene, Biotechnol. Tech. 11, 375-378).

Äußerst bevorzugt ist die Immobilisierung an Eupergit®, insbesondere Eupergit C® und Eupergit 250L® (Röhm)

5 (Eupergit.RTM. C, a carrier for immobilization of enzymes of industrial potential. Katchalski-Katzir, E.; Kraemer, D. M. Journal of Molecular Catalysis B: Enzymatic (2000), 10(1-3), 157-176.)

Gleichfalls bevorzugt ist die Immobilisierung an Ni-NTA in 10 Kombination mit dem His-Tag (Hexa-Histidin) ergänzten Polypeptid (Purification of proteins using polyhistidine affinity tags. Bornhorst, Joshua A.; Falke, Joseph J. Methods in Enzymology (2000), 326, 245-254). Die Verwendung als CLECs ist ebenfalls denkbar (St. Clair, N.; Wang, Y.- 15 F.; Margolin, A. L. (2000), Cofactor-bound cross-linked enzyme crystals (CLEC) of alcohol dehydrogenase, Angew. Chem. Int. Ed. 39, 380-383).

Durch diese Maßnahmen kann es gelingen aus Polypeptiden, welche durch organische Solventien instabil werden, solche 20 zu generieren, die in Gemischen von wässrigen und organischen Lösungsmitteln bzw. ganz in Organik stabil sind und arbeiten können.

Ganzzellkatalysatoren werden im Allgemeinen in Form freier oder immobilisierter Zellen eingesetzt. Hierzu wird die aktive Zellmasse in einer hydantoinhaltigen Lösung resuspendiert. Die Zellkonzentration beträgt dabei zwischen 1-100g/l. Die Konzentration des Hydantoins liegt zwischen 0,1 und 2 molar. Als Lösungsmittel wird bevorzugt H₂O verwendet, wobei jedoch auch Mischungen von organischen 30 Lösungsmitteln und H₂O einsetzbar sind. Der pH-Wert wird entweder nicht geregelt oder mittels gängiger Puffer bzw. durch kontinuierliche pH-Statisierung zwischen pH6 und pH10 konstant gehalten. Die Reaktionstemperatur liegt typischerweise zwischen 20°C und 90°C. In Abhängigkeit der 35 verwendeten Hydantoinase werden zweiwertige Metall-Ionen in Konzentrationen von 0,1-5mM hinzugesetzt. Bevorzugte

Metallionen sind dabei Mn^{2+} , Zn^{2+} oder Co^{2+} .

In Bezug auf den Einsatz der einzelnen Enzyme kann in äquivalenter Art und Weise verfahren werden.

Die durch den Einsatz der erfindungsgemäßen

- 5 Hydantoinrazematen in wie z.B. oben beschriebener Weise hergestellten Produkte werden nach gängigen Verfahren aufgearbeitet. Vorteilhaft ist jedoch die Aufarbeitung durch Ionenaustauschchromatographie. Hierdurch wird das Produkt vom bei der Reaktion entstehenden Salzen befreit.
- 10 Das Eluat wird ggf. mit Aktivkohle geklärt und die entstandene enantiomerenangereicherte Aminosäure oder N-Carbamoyl-Aminosäure durch Einengung des Lösungsmittels ausgefällt und getrocknet.

Die Kopplung einer enzymatischen Razemisierung mit einer

- 15 enantioselektiven Hydrolyse zum Screenen von Hydantoinrazemaseaktivitäten wurde bisher nicht zur Erzeugung verbesserter Hydantoinrazematen angewendet. Für eine besonders erfolgreiche Anwendung des erfindungsgemäßen Verfahrens sollten mehrere Voraussetzungen erfüllt sein:

- 20 1. Die chemische Razemisierungsgeschwindigkeit des im Screening verwendeten enantiomerenreinen Hydantoins muss sehr viel kleiner sein, als die Geschwindigkeit der enzymatisch katalysierten Reaktion.
- 25 2. Die enantioselektive enzymatische Hydrolyse mittels der Hydantoinase muss sehr viel schneller erfolgen als die enzymatische Razemisierung des Hydantoins.

Für aliphatisch substituierte Hydantoine ist, durch deren langsame chemische Razemisierung bedingt, Punkt 1 gegeben.

Punkt 2 kann durch eine gezielte Auswahl von geeigneten

- 30 Hydantoinasen (s. weiter vorne) erfüllt werden.

Mit den Aussagen des Standes der Technik wird die vorliegende Erfindung nicht nahegelegt, da diesem keinerlei

Hinweise auf die weiter oben genannten Voraussetzungen zu entnehmen sind.

Sämtliche der gezeigten Mutanten weisen an der Aminosäureposition 79 eine Mutation auf, was die Bedeutung

5 dieser Position für die Enzymfunktion erstmalig aufzeigt.

Interessant ist, dass die Aminosäuren, welche diese

Position umgeben, für sämtliche bekannten

Hydantoinrazematen vollständig konserviert sind. Hieraus ergibt sich, dass für andere Hydantoinrazematen welche das

10 oben beschriebene Sequenzmotiv enthalten und eine hohe

Homologie (>40% Sequenzidentität) aufweisen durch

ortsspezifische Mutagenese an Pos. 79 verbesserte

Enzymvarianten erzeugt werden können, was bisher aus dem Stand der Technik nicht herleitbar war.

15 Unter optisch angereicherten (enantiomerenangereicherten; enantiomer angereicherten) Verbindungen wird im Rahmen der Erfindung das Vorliegen einer optischen Antipode im Gemisch mit der anderen in >50 mol-% verstanden.

Unter dem Begriff Nukleinsäuresequenzen werden alle Arten

20 von einzelsträngiger oder doppelsträngiger DNA als auch RNA oder Gemische derselben subsumiert.

Die Verbesserung der Aktivität und/oder Selektivität und/oder Stabilität bedeutet erfindungsgemäß, dass die Polypeptide aktiver und/oder selektiver bzw. weniger

25 selektiv oder unter den verwendeten Reaktionsbedingungen stabiler sind. Während die Aktivität und die Stabilität der Enzyme für die technische Anwendung naturgemäß möglichst hoch sein sollte, ist in Bezug auf die Selektivität dann von einer Verbesserung die Rede, wenn entweder die

30 Substratselektivität abnimmt, die Enantioselektivität der Enzyme jedoch gesteigert ist.

Von den beanspruchten Polypeptiden und den

Nukleinsäuresequenzen werden erfindungsgemäß auch solche

Sequenzen umfaßt, die eine Homologie (exclusive der

natürlichen Degeneration) größer als 70% (in Bezug auf die Nukleinsäuresequenz) bzw. > 40% oder 80% (in Bezug auf die Polypeptide), bevorzugt größer als 90%, 91%, 92%, 93% oder 94%, mehr bevorzugt größer als 95% oder 96% und besonders bevorzugt größer als 97%, 98% oder 99% zu einer dieser Sequenzen aufweisen, sofern die Wirkungsweise bzw. Zweck einer solchen Sequenz erhalten bleibt. Der Ausdruck "Homologie" (oder Identität) wie hierin verwendet, kann durch die Gleichung $H (\%) = [1 - V/X] \times 100$ definiert werden, worin H Homologie bedeutet, X die Gesamtzahl an Nukleobasen/Aminosäuren der Vergleichssequenz ist und V die Anzahl an unterschiedlichen Nukleobasen/Aminosäuren der zu betrachtenden Sequenz bezogen auf die Vergleichssequenz ist. Auf jeden Fall sind mit dem Begriff Nukleinsäuresequenzen, welche für Polypeptide codieren, alle Sequenzen umfaßt, die nach Maßgabe der Degeneration des genetischen Codes möglich erscheinen.

Der Ausdruck "unter stringenten Bedingungen" wird hierin wie bei Sambrook et al. (Sambrook, J.; Fritsch, E. F. und 20 Maniatis, T. (1989), Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, New York) beschrieben, verstanden. Bevorzugt liegt eine stringenten Hybridisierung gemäß der vorliegenden Erfindung vor, wenn nach Waschen für eine Stunde mit 1 x SSC (150 mM Natriumchlorid, 15 mM Natriumcitrat, pH 7.0) und 0,1 % SDS (Natriumdodecylsulfat) bei 50 °C, bevorzugt bei 55 °C, mehr bevorzugt bei 62 °C und am meisten bevorzugt bei 68 °C und mehr bevorzugt für 1 Stunde mit 0,2 x SSC und 0,1 % SDS bei 50 °C, bevorzugter bei 55 °C, mehr bevorzugt bei 62 °C und am 30 meisten bevorzugt bei 68 °C noch ein positives Hybridisierungssignal beobachtet wird.

Die in dieser Schrift genannten Literaturstellen gelten als von der Offenbarung mitumfaßt.

Der Organismus *Arthrobacter aurescens* DSM3747 wurde durch 35 die Rütgerswerke Aktiengesellschaft am 28.05.86 bei der

030115 AM / IP

21

Deutschen Sammlung für Mikroorganismen GmbH, Mascheroder
Weg 1b, 38124 Braunschweig hinterlegt.

Beispiele

**Beispiel 1: Erzeugung Hydantoinrazemasemutanten -
Zufallsmutagenese**

5 0,25ng des Vektors pOM21 (Plasmidkarte siehe Fig.1; Sequenz
siehe Seq.ID.Nr.13) (PCT/US00/08159) wurde als Template in
einem 100 μ l PCR Reaktionsmix bestehend aus PCR-Puffer (10
mM Tris, 1.5 mM MgCl₂, 50 mM KC1, pH 8.5), 200 μ M dTTP, 200
 μ M dGTP, 200 μ M dATP, 200 μ M dCTP, 50 pmol des jeweiligen
10 Primers (siehe Seq.ID.Nr.11 und 12) und 2,5 U Taq-
Polymerase (Roche) eingesetzt. Nach 30 Zyklen wurde das
Amplifikat mittels Gelextraktion (QiaexII Gel-
Extraktionskit) aufgereinigt und in den Vektor pOM21
mittels den Restriktionsenzymen NdeI und PstI subkloniert.
15 Das Ligationsprodukt wurde zur Transformation von
hydantoinasepositiver Stämme verwendet (siehe Beispiel 2).

**Beispiel 2: Herstellung von hydantoinasepositiven Stämmen
und einer Mutantenbank**

Chemisch kompetente *E.coli* JM109 (z.B. von Promega) wurden
20 mit 10ng des Plasmids pDHYD (siehe Fig.2; siehe Seq.ID.Nr.
15) (**Herstellung?**) transformiert, welches das D-
Hydantoinasegen aus *Arthrobacter crystallopoietes* DSM20117
unter Kontrolle eines Rhamnose-Promotors trägt. Die
vollständige Sequenz des Plasmids ist in Seq.ID.Nr. 15
angegeben. Der so erzeugte hydantoinasepositive Stamm wurde
wiederum chemisch kompetent gemacht und zur Herstellung der
Mutantenbank mit dem Ligationsprodukt der
Hydantoinrazemase-Zufallsmutagenese aus Beispiel 1
transformiert. Die auf Ampicillin- und Chloramphenicol-
haltigen Agarplatten ausgestrichenen Kolonien der
30 Mutantenbank wurden anschliessend einem Screening
unterworfen, welches in Beispiel 3 beschrieben wird.

Beispiel 3: Screening nach Hydantoinrazemasesmutanten mit verbesserten Enzymeigenschaften

Einzelne Kolonien der Mutantenbank wurden in 96-Well-Platten überimpft, welche mit 100µl pro Well Rhamnose

5 (2g/l) und ZnCl₂ (1mM) supplementiertem LB-Medium (5g/l Hefeextrakt, 10g/l Trypton, 10g/l NaCl) gefüllt waren. Die Platten wurden für 20 Stunden bei 30°C inkubiert.

Anschliessen wurden 100µl Screening-Substrat (100mM L-Ethylhydantoin, 50mg/l Cresol Rot, pH 8.5) zu jedem Well zugegeben und die Platten für 4 Stunden bei 20°C inkubiert. Wells mit verbesserten Hydantoinrazemasesmutanten konnten durch eine intensivere Gelbfärbung im Vergleich zum Wildtyp direkt per Auge, oder unter Verwendung eines Spektralphotometers bei 580nm identifiziert werden.

15 Beispiel 4: Charakterisierung von Hydantoinrazemasesmutanten mit verbesserten Enzymeigenschaften

Die im Screening identifizierten Razemasesmutanten wurden anschliessend mittels HPLC-Analyse auf ihre Aktivität im Vergleich zum Wildtyp untersucht und die entsprechenden

20 Mutationen mittels Sequenzierung bestimmt. Hierzu wurde von einzelnen Kolonien der unterschiedlichen Klone Plasmide isoliert (Qiagen Mini-Prep Kit) und sequenziert. Die selben Klone wurden zur Herstellung aktiver Biomasse verwendet.

Eine Übernachtkultur (OD₆₀₀=4) der jeweiligen Klone wurde hierzu 1:100 in 100ml Rhamnose (2g/l) und ZnCl₂ (1mM) supplementiertem LB-Medium (5g/l Hefeextrakt, 10g/l Trypton, 10g/l NaCl) verdünnt und 18 Stunden bei 30°C und 250UPM inkubiert. Die Biomasse wurde mittels Zentrifugation (10min, 10.000g) pelletiert und der Überstand verworfen. 2g

30 aktive Biomasse wurde anschliessend in 50ml der Substratlösung (100mM L-Ethylhydantoin, pH 8.5) resuspendiert und bei 37°C inkubiert. Nach verschiedenen Zeiten wurden Proben genommen, die Biomasse durch Zentrifugation (5min, 13.000 UPM) abgetrennt und der

Überstand mittels HPLC auf die Konzentration der entstandenen N-Carbamoyl-aminobuttersäure analysiert.

Beispiel 5 Herstellung von L-Aminosäuren unter Verwendung verbesserter Hydantoinrazematen

5 Ein mit pOM21-BB5 und pOM22 Fig. 3 (siehe Seq.ID.Nr.14) (PCT/US00/08159) transformierter Stamm von *E.coli* JM109 wurde bei 30°C in Ampicillin (100µg/l) und Chloramphenicol (50µg/l) -haltigem sowie mit 2g/l Rhamnose versetztem LB-Medium für 18 Stunden unter Schütteln (250 U/min)

10 inkubiert. Die Biomasse wurde durch Zentrifugation pelletiert und mit einem entsprechenden Volumen von 100mM DL-Ethlyhydantoinlösung, pH 8,5 und 1mM CoCl₂ so resuspendiert, dass sich eine Zellkonzentration von 30g/l ergibt. Diese Reaktionslösung wurde für 10 Stunden bei 37°C

15 inkubiert. Anschliessend wurden die Zellen durch Zentrifugation (30 min, 5000g) abgetrennt und der klare Überstand mittels HPLC auf die entstandene Aminosäure analysiert. Zur Aufarbeitung der entstandenen Aminosäure wurde das Volumen des Überstandes auf die Hälfte reduziert

20 und 1:2 mit Methanol versetzt. Die ausgefällte Aminosäure wurde anschliessend filtriert und getrocknet. Die Gesamtausbeute der Aminosäure betrug >60%.

Beispiel 6 Herstellung von D-Aminosäuren unter Verwendung verbesserter Hydantoinrazematen

25 Ein mit pOM21-BB5 und pJAVIER16 Fig. 4 (siehe Seq.ID.Nr.16) (**Herstellung?**) transformierter Stamm von *E.coli* JM109 wurde bei 30°C in Ampicillin (100µg/l) und Chloramphenicol (50µg/l) -haltigem sowie mit 2g/l Rhamnose versetztem LB-Medium für 18 Stunden unter Schütteln (250 U/min)

30 inkubiert. Die Biomasse wurde durch Zentrifugation pelletiert und mit einem entsprechenden Volumen von 100mM DL-Ethlyhydantoinlösung, pH 8,5 und 1mM CoCl₂ so resuspendiert, dass sich eine Zellkonzentration von 30g/l ergibt. Diese Reaktionslösung wurde für 10 Stunden bei 37°C

inkubiert. Anschliessend wurden die Zellen durch Zentrifugation (30 min, 5000g) abgetrennt und der klare Überstand mittels HPLC auf die entstandene Aminosäure analysiert. Zur Aufarbeitung der entstandenen Aminosäure 5 wurde das Volumen des Überstandes auf die Hälfte reduziert und 1:2 mit Methanol versetzt. Die ausgefällte Aminosäure wurde anschliessend filtriert und getrocknet. Die Gesamtausbeute der Aminosäure betrug >60%.

Patentansprüche:

1. Screeningverfahren für Hydantoinrazematen,
dadurch gekennzeichnet, dass man
 - a) eine enantioselektive Hydantoinase und
 - 5 b) die zu prüfende Hydantoinrazemase, welche eine verglichen mit der Hydantoinase unter a) langsamere Umsetzungsrate aufweist, auf
 - c) ein chirales Hydantoin einwirken
 - 10 lässt, welches in zur Selektivität der Hydantoinase entgegengesetzter enantiomerenangereicherten Form eingesetzt wird, und
 - d) die resultierende N-Carbamoyl-Aminosäure oder die freigesetzten Protonen zeitabhängig detektiert.
2. Verfahren nach Anspruch 1,
dadurch gekennzeichnet, dass man
15 ein aliphatisch substituiertes Hydantoin einsetzt.
3. Verfahren nach einem oder mehreren der vorhergehenden
Ansprüche,
dadurch gekennzeichnet, dass man
20 ein Hydantoinase aus *Arthrobacter crystallopoietes*
einsetzt.
4. Verfahren nach einem oder mehreren der vorhergehenden
Ansprüche,
dadurch gekennzeichnet, dass
25 das Verhältnis der Geschwindigkeitskonstanten der
Hydantoinase zur Hydantoinrazemase (k_{Hyd}/k_{Raz}) > 2 ist.
5. Verfahren zur Herstellung von verbesserten
Hydantoinrazematen,
dadurch gekennzeichnet, dass man
 - 30 a) die Nukleinsäuresequenz codierend für die Hydantoinrazemase einer Mutagenese unterwirft,
 - b) die aus a) erhältlichen Nukleinsäuresequenzen in einen geeigneten Vektor kloniert und diesen in ein

geeignetes Expressionssystem transferiert und
c) die gebildeten Hydantoinrazematen mit verbesserter
Aktivität und/oder Selektivität und/oder Stabilität
mittels eines Verfahrens nach einem oder mehreren der
5 Ansprüche 1 bis 4 detektiert und isoliert.

6. rec-Polypeptide oder diese codierende
Nukleinsäuresequenzen erhältlich nach Anspruch 5.
7. Verwendung der Polypeptide gemäß Anspruch 6 zur
Herstellung von enantiomerenangereicherten N-
10 Carbamoyl-Aminosäure oder Aminosäuren.
8. Verwendung der Nukleinsäuresequenzen gemäß 6 zur
Herstellung von Ganzzellkatalysatoren.
9. Hydantoinrazemase aufweisend in Position 79 einen
Aminosäureaustausch mit einer Aminosäure ausgewählt
15 aus der Gruppe bestehend aus A, R, N, D, C, Q, E, H,
I, L, K, M, F, P, S, T, Y oder V.
10. Hydantoinrazematen aufweisend die Konsensussequenz
FX₁DX₂GL (Seq. 1), wobei X₂ P oder T darstellt und X₁
in der Position 79 eine Aminosäure ausgewählt aus der
20 Gruppe A, R, N, D, C, Q, E, H, I, L, K, M, F, P, S, T,
Y oder V darstellt darstellt.
11. Isolierte Nukleinsäuresequenz codierend für eine
Hydantoinrazemase ausgewählt aus der Gruppe:
 - a) einer Nukleinsäuresequenz codierend für eine
Hydantoinrazemase gemäß Anspruch 9 und/oder 10,
 - b) einer Nukleinsäuresequenz, die unter stringenten
Bedingungen mit der Nukleinsäuresequenz codierend
für eine Hydantoinrazemase gemäß Anspruch 9
und/oder 10 oder der dazu komplementären Sequenz
30 hybridisiert,
 - c) einer Nukleinsäuresequenz gemäß den Seq.ID.Nr. 3,
5, 7 oder 9 oder solchen mit einer Homologie von
> 80% zu diesen,

- d) einer Nukleinsäuresequenz aufweisend 15 aufeinanderfolgende Nukleotide der Sequenzen Seq.ID.Nr. 3, 5, 7 oder 9..
- 12. Ganzzellkatalysator aufweisend ein kloniertes Gen für eine Hydantoinrazemase gemäß den Ansprüchen 9 und/oder 10.
- 13. Plasmide, Vektoren oder Mikroorganismen aufweisend eine Nukleinsäuresequenz gemäß Anspruch 9 und/oder 10.
- 14. Primer zur Herstellung der Nukleinsäuresequenzen nach Anspruch 9 und/oder 10 mittels PCR.

Zusammenfassung:

Die vorliegende Erfindung bezieht sich auf ein Screeningverfahren für Hydantoinrazemases und neue Hydantoinrazemases, die sie codierenden
5 Nukleinsäuresequenzen und ein Verfahren zur Mutagenese.

Hydantoinrazemases sind im Zusammenhang mit der Erzeugung von enantiomerenangereicherten Aminosäuren aus racemischen Hydantoinen von Interesse.

Abb. 1:

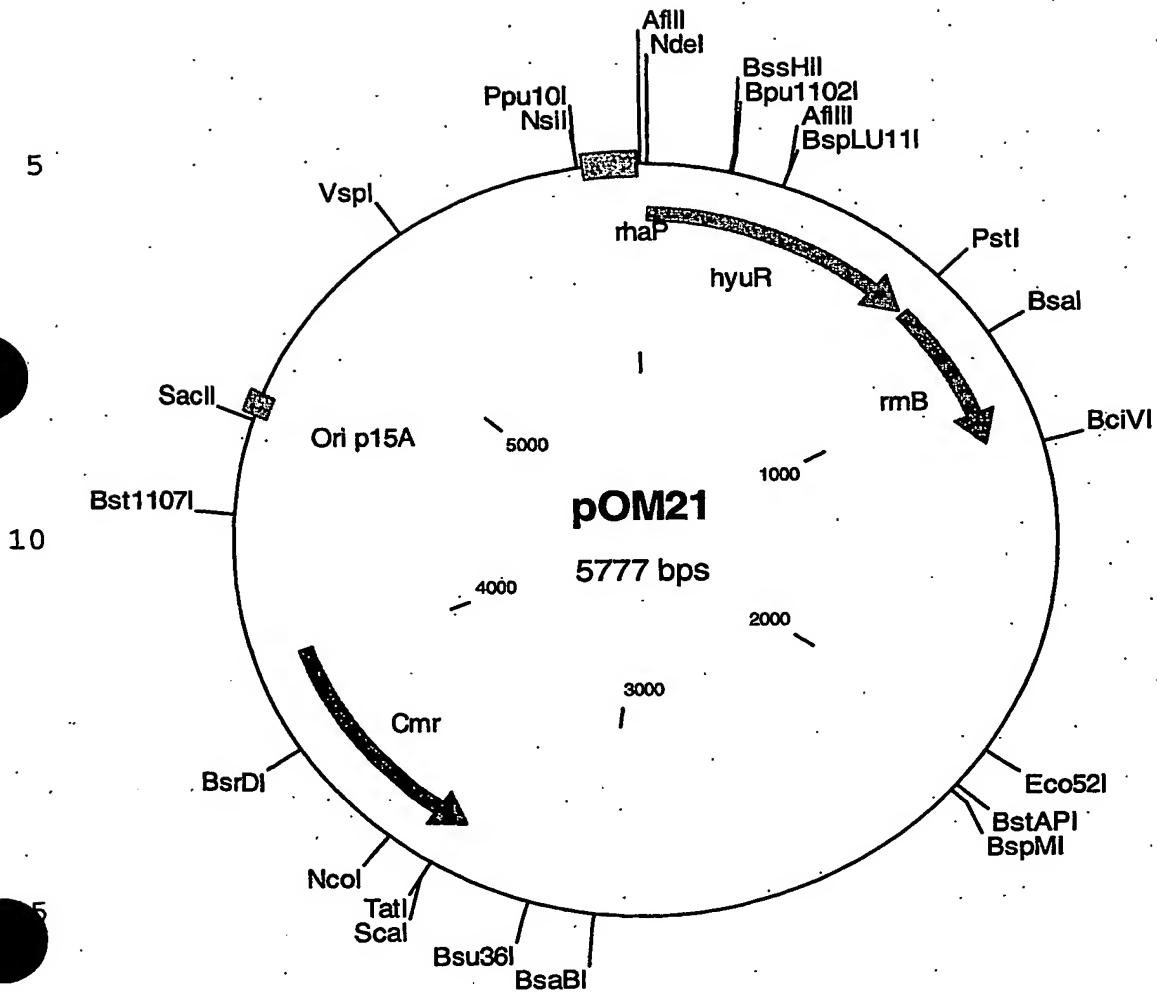


Abb. 2:

5

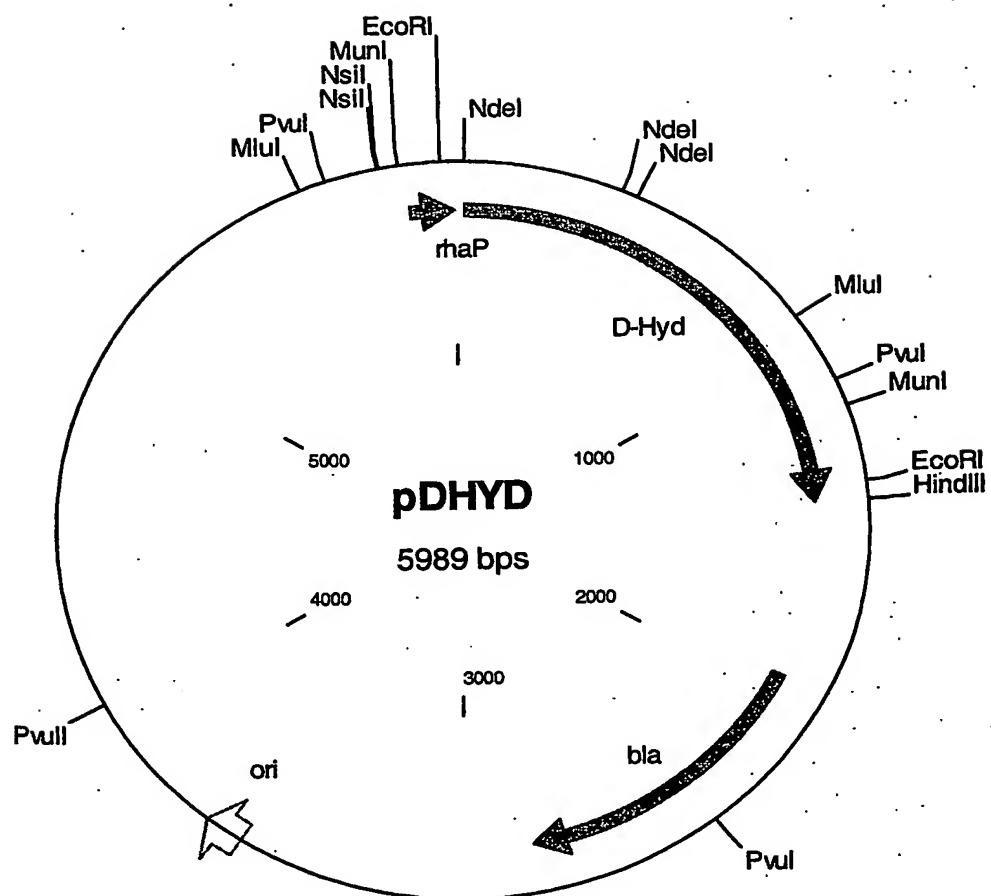


Fig: 3

5

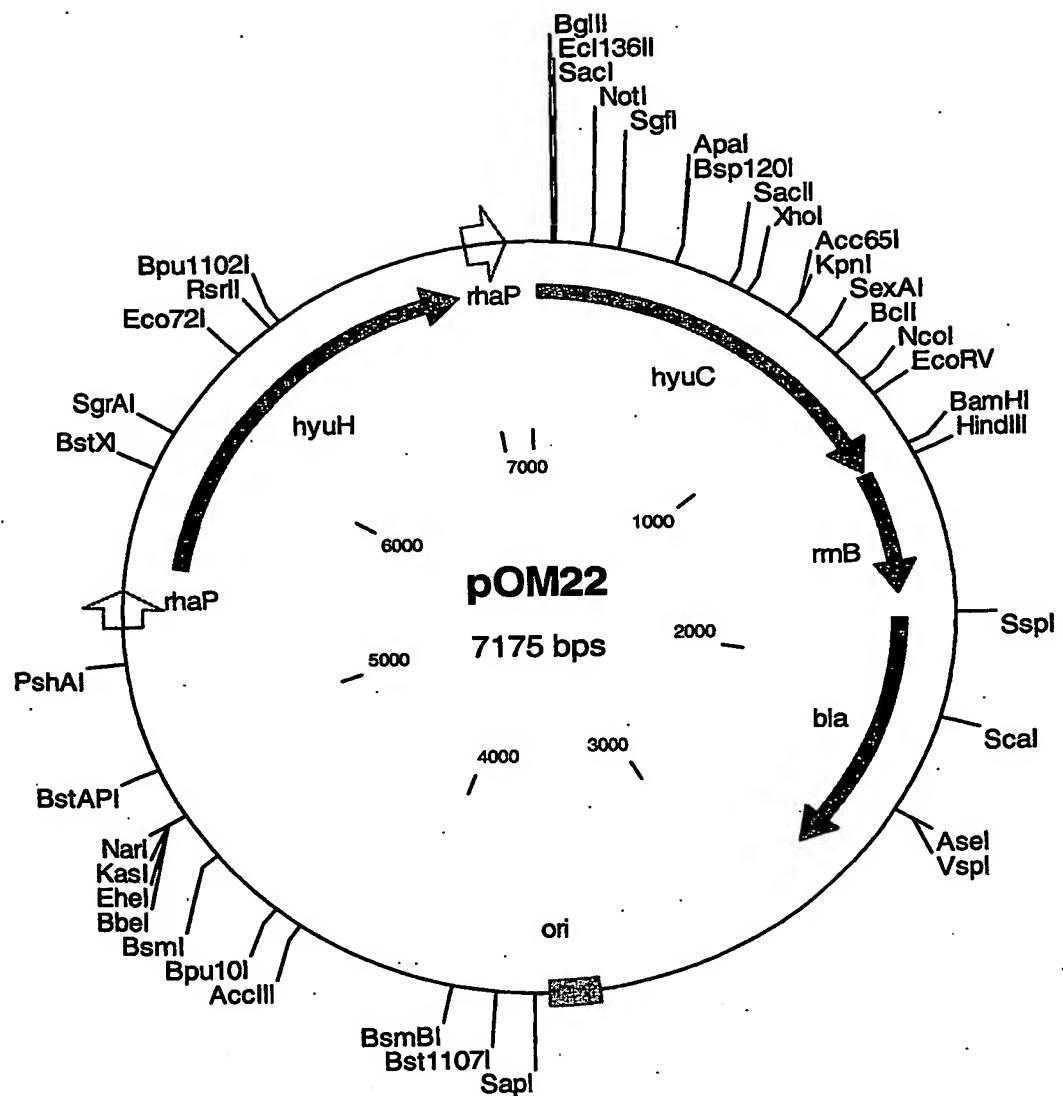
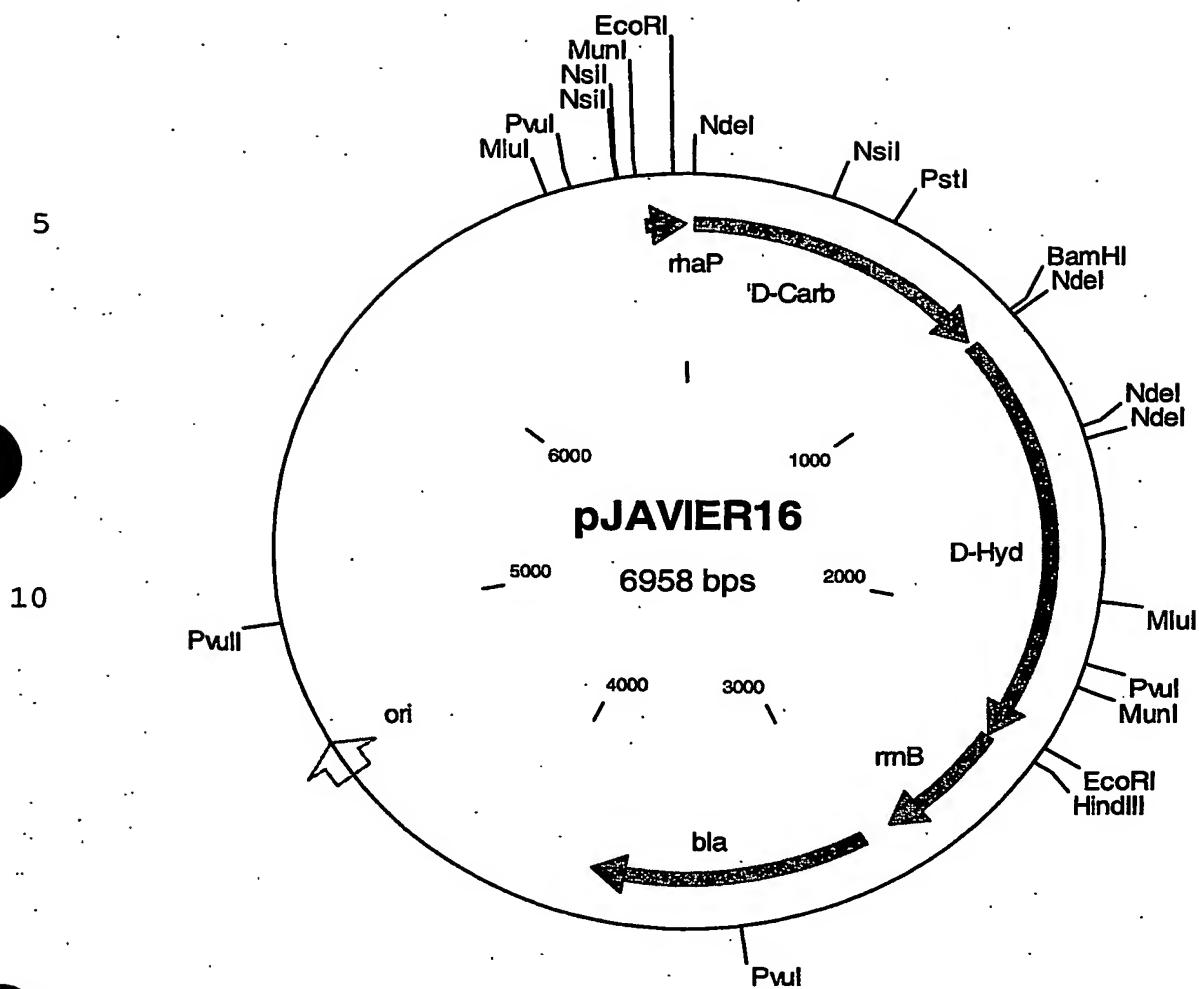


Fig. 4



030115 AM / IP

1

SEQUENZPROTOKOLL

<110> Degussa AG

5 <120> Screeningverfahren für Hydantoinrazematen

<130> 030115 AM / IP

<140>

10 <141>

<160> 16

15 <170> PatentIn Ver. 2.1

<210> 1

<211> 6

<212> PRT

<213> Künstliche Sequenz

0 <220>

<223> Beschreibung der künstlichen Sequenz:
Konsensussequenz

25 <400> 1

Phe Xaa Asp Xaa Gly Leu
1 5

30 <210> 2

<211> 236

<212> PRT

<213> Arthrobacter crystallopoietes

35 <400> 2

Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
1 5 10 15

Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
20 25 30

Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
35 40 45

45 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
50 55 60

Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Gly Asp
65 70 75 80

50 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
85 90 95

55 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
100 105 110

Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
115 120 125

Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

5 Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

10 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

15 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

20 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

<210> 3
 <211> 711
 25 <212> DNA
 <213> Künstliche Sequenz

<220>
 30 <223> Beschreibung der künstlichen Sequenz: 1BG7

<220>
 <221> CDS
 <222> (1)...(711)

35 <400> 3
 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ala Leu Thr Glu
 1 5 10 15

tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att 96
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

45 tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt 144
 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct 192
 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 50 55 60

gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc agg gat 240
 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Arg Asp
 65 70 75 80

55 ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga 288
 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

gtt	gcc	gaa	gct	gca	atc	cac	atg	tct	tca	ttc	gtc	gcg	gcc	acc	ttc	336	
Val	Ala	Glu	Ala	Ala	Ile	His	Met	Ser	Ser	Phe	Val	Ala	Ala	Thr	Phe		
100								105						110			
5	tcc	att	gtc	agc	atc	ctc	ccg	agg	gtc	agg	aaa	cat	ctg	cac	gaa	ctg	384
	Ser	Ile	Val	Ser	Ile	Leu	Pro	Arg	Val	Arg	Lys	His	Leu	His	Glu	Leu	
	115							120					125				
10	gta	cg	caa	g	cg	gg	g	cg	aa	t	cg	ct	g	cc	aa	ct	432
	Val	Arg	Gln	Ala	Gly	Ala	Thr	Asn	Arg	Leu	Ala	Ser	Ile	Lys	Leu	Pro	
	130							135					140				
15	aat	ct	tg	gg	gt	at	g	tc	at	g	g	aa	cat	g	ca	ct	480
	Asn	Leu	Gly	Val	Met	Ala	Phe	His	Glu	Asp	Glu	His	Ala	Ala	Leu	Glu	
	145							150					155			160	
	acg	ct	aa	ca	g	cc	g	cc	aa	g	g	g	tc	g	cc	g	528
	Thr	Leu	Lys	Gln	Ala	Ala	Lys	Glu	Ala	Val	Gln	Glu	Asp	Gly	Ala	Glu	
								165					170			175	
0	tcg	at	tg	ct	gg	tg	cc	gg	at	tg	gg	tt	gg	cg	ct	576	
	Ser	Ile	Val	Leu	Gly	Cys	Ala	Gly	Met	Val	Gly	Phe	Ala	Arg	Gln	Leu	
	180							185					190				
25	agc	gac	gaa	ct	gg	gt	cc	ct	at	gac	cc	gt	gag	gca	g	tc	624
	Ser	Asp	Glu	Leu	Gly	Val	Pro	Val	Ile	Asp	Pro	Val	Glu	Ala	Ala	Cys	
	195							200					205				
30	cg	gt	gg	g	ag	tt	g	tc	g	ct	gg	tc	ac	ag	aa	g	672
	Arg	Val	Ala	Glu	Ser	Leu	Val	Ala	Leu	Gly	Tyr	Gln	Thr	Ser	Lys	Ala	
	210							215					220				
35	aa	tc	tg	ta	aa	cc	ac	g	ag	aa	g	ag	ta	ct	ta	711	
	Asn	Ser	Tyr	Gln	Lys	Pro	Thr	Glu	Lys	Gln	Tyr	Leu					
	225							230					235				
	<210> 4																
	<211> 237																
	<212> PRT																
	<213> Künstliche Sequenz																
	<223> Beschreibung der künstlichen Sequenz:1BG7																
	<400> 4																
45	Met	Arg	Ile	Leu	Val	Ile	Asn	Pro	Asn	Ser	Ser	Ser	Ala	Leu	Thr	Glu	
	1				5					10				15			
	Ser	Val	Ala	Asp	Ala	Ala	Gln	Gln	Val	Val	Ala	Thr	Gly	Thr	Ile	Ile	
									20				25		30		
50	Ser	Ala	Ile	Asn	Pro	Ser	Arg	Gly	Pro	Ala	Val	Ile	Glu	Gly	Ser	Phe	
	35								40					45			
	Asp	Glu	Ala	Leu	Ala	Thr	Phe	His	Leu	Ile	Glu	Glu	Val	Glu	Arg	Ala	
55									55					60			
	Glu	Arg	Glu	Asn	Pro	Pro	Asp	Ala	Tyr	Val	Ile	Ala	Cys	Phe	Arg	Asp	
	65								70					75		80	

Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 5 100 105 110

Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

10 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

15 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

20 Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

25 Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

30 <210> 5
 <211> 711
 <212> DNA
 35 <213> Künstliche Sequenz
 <220>
 <223> Beschreibung der künstlichen Sequenz: 3CH11
 <220>
 <221> CDS
 <222> (1)...(711)
 <400> 5
 45 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ala Leu Thr Glu
 1 5 10 15

50 tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att 96
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

55 tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt 144
 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct 192
 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

5	gag	cg	gaa	aac	ccg	ccc	gac	gcc	tac	gtc	atc	gca	tgt	ttc	gag	gat	240
	Glu	Arg	Glu	Asn	Pro	Pro	Asp	Ala	Tyr	Val	Ile	Ala	Cys	Phe	Glu	Asp	
	65				70.						75					80	
10	ccg	gga	ctt	gac	g	cg	gt	a	ag	ct	g	ac	ag	cc	gt	gt	288
	Pro	Gly	Leu	Asp	Ala	Val	Lys	Glu	Leu	Thr	Asp	Arg	Pro	Val	Val	Gly	
	85					90				95							
15	gtt	g	cc	gaa	g	c	at	c	ac	at	t	c	t	ca	tc	g	336
	Val	Ala	Glu	Ala	Ala	Ile	His	Met	Ser	Ser	Phe	Val	Ala	Ala	Thr	Phe	
	100					105					110						
20	tcc	att	gt	tc	ag	c	at	c	tc	cc	ag	aa	ca	ct	g	384	
	Ser	Ile	Val	Ser	Ile	Leu	Pro	Arg	Val	Arg	Lys	His	Leu	His	Glu	Leu	
	115					120					125						
25	gta	cg	ca	g	cg	gg	g	cg	aa	ca	cg	cc	tc	at	cc	432	
	Val	Arg	Gln	Ala	Gly	Ala	Thr	Asn	Arg	Leu	Ala	Ser	Ile	Lys	Leu	Pro	
	130					135					140						
30	aat	ct	g	gg	gt	at	g	cc	t	tc	at	aa	ct	cc	ca	480	
	Asn	Leu	Gly	Val	Met	Ala	Phe	His	Glu	Asp	Glu	His	Ala	Ala	Leu	Glu	
	145					150					155					160	
35	ac	ct	aa	ca	g	cc	g	ag	g	cg	ca	gg	cc	g	ag	528	
	Thr	Leu	Lys	Gln	Ala	Ala	Lys	Glu	Ala	Val	Gln	Glu	Asp	Gly	Ala	Glu	
	165						165			170					175		
40	tcg	at	gt	ct	gg	tgc	g	cc	at	gt	gg	tt	g	cg	ct	576	
	Ser	Ile	Val	Leu	Gly	Cys	Ala	Gly	Met	Val	Gly	Phe	Ala	Arg	Gln	Leu	
	180						180			185					190		
45	ag	g	ac	g	aa	ct	gg	gt	c	at	g	cc	gt	ca	gt	624	
	Ser	Asp	Glu	Leu	Gly	Val	Pro	Val	Ile	Asp	Pro	Val	Glu	Ala	Ala	Cys	
	195						195			200					205		
50	cg	gt	g	cc	g	ag	gt	gt	ct	g	cc	g	ac	gg	aa	672	
	Arg	Val	Ala	Glu	Ser	Leu	Val	Ala	Leu	Gly	Tyr	Gln	Thr	Ser	Lys	Ala	
	210					215					220						
55	aa	tc	ta	ca	aa	cc	ac	g	ag	ca	g	cc	g	ac	gg	711	
	Asn	Ser	Tyr	Gln	Lys	Pro	Thr	Glu	Lys	Gln	Tyr	Leu					
	225					230					235						
60	<210>	6															
	<211>	237															
	<212>	PRT															
65	<213>	Künstliche Sequenz															
	<223>	Beschreibung der künstlichen Sequenz:3CH11															
70	<400>	6															
	Met	Arg	Ile	Leu	Val	Ile	Asn	Pro	Asn	Ser	Ser	Ser	Ala	Leu	Thr	Glu	
	1					5				10					15		
75	Ser	Val	Ala	Asp	Ala	Ala	Gln	Gln	Val	Val	Ala	Thr	Gly	Thr	Ile	Ile	
	20						20			25					30		

Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

5 Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Glu Asp
 65 70 75 80

10 Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110

15 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

20 Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

25 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

30 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

45 <210> 7
 <211> 711
 <212> DNA
 <213> Künstliche Sequenz

50 <220>
 <223> Beschreibung der künstlichen Sequenz:AE3

55 <220>
 <221> CDS
 <222> (1) .. (711)

<400> 7
 atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa 48
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ala Leu Thr Glu
 1 5 10 15

tcg gtt gcg gac gca gca caa gtt gtc gcg acc ggc acc ata att 96
 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile

	20	25	30	
5	tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe 35 40 45			144
10	gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala 50 55 60			192
15	gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc cag gat Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Gln Asp 65 70 75 80			240
20	ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly 85 90 95			288
25	gtt gcc gaa gct gca atc cac atg tct tca ttc gtc gcg gcc acc ttc Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe 100 105 110			336
30	tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu 115 120 125			384
35	gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro 130 135 140			432
40	aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu 145 150 155 160			480
45	acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu 165 170 175			528
50	tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu 180 185 190			576
55	agc gac gaa ctc ggc gtc cct gtc atc gac ccc gtc gag gca gct tgc Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys 195 200 205			624
60	cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala 210 215 220			672
65	aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu 225 230 235			711
70	<210> 8 <211> 237 <212> PRT <213> Künstliche Sequenz			

<223> Beschreibung der künstlichen Sequenz:AE3

<400> 8

5	Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu	15
	1 5 10	
10	Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile	30
	20 25 30	
15	Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe	45
	35 40 45	
20	Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala	60
	50 55 60	
25	Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Gln Asp	80
	65 70 75 80	
30	Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly	95
	85 90 95	
35	Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe	110
	100 105 110	
40	Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu	125
	115 120 125	
45	Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro	140
	130 135 140	
50	Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu	160
	145 150 155 160	
55	Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu	175
	165 170 175	
60	Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu	190
	180 185 190	
65	Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys	205
	195 200 205	
70	Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala	220
	210 215 220	
75	Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu	235
	225 230 235	
80	<210> 9	
	<211> 711	
	<212> DNA	
	<213> Künstliche Sequenz	
85	<220>	
	<223> Beschreibung der künstlichen Sequenz:BB5	
	<220>	
	<221> CDS	

<222> (1)..(711)

<400> 9

5	atg aga atc ctc gtg atc aac ccc aac agt tcc agc gcc ctt act gaa Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu 1 5 10 15	48
10	tcg gtt gcg gac gca gca caa caa gtt gtc gcg acc ggc acc ata att Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile 20 25 30	96
15	tct gcc atc aac ccc tcc aga gga ccc gcc gtc att gaa ggc agc ttt Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe 35 40 45	144
20	gac gaa gca ctg gcc acg ttc cat ctc att gaa gag gtg gag cgc gct Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala 50 55 60	192
25	gag cgg gaa aac ccg ccc gac gcc tac gtc atc gca tgt ttc ttg gat Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Leu Asp 65 70 75 80	240
30	ccg gga ctt gac gcg gtc aag gag ctg act gac agg cca gtg gta gga Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly 85 90 95	288
35	gtt gcc gaa gct gca atc cac atg tct tca ttc gtc gcg gcc acc ttc Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe 100 105 110	336
40	tcc att gtc agc atc ctc ccg agg gtc agg aaa cat ctg cac gaa ctg Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu 115 120 125	384
45	gta cgg caa gcg ggg gcg acg aat cgc ctc gcc tcc atc aag ctc cca Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro 130 135 140	432
50	aat ctg ggg gtg atg gcc ttc cat gag gac gaa cat gcc gca ctg gag Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu 145 150 155 160	480
55	acg ctc aaa caa gcc gcc aag gag gcg gtc cag gag gac ggc gcc gag Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu 165 170 175	528
60	tcg ata gtg ctc gga tgc gcc ggc atg gtg ggg ttt gcg cgt caa ctg Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu 180 185 190	576
65	agc gac gaa ctc ggc gtc cct gtc atc gac ccc gtc gag gca gct tgc Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys 195 200 205	624
70	cgc gtg gcc gag agt ttg gtc gct ctg ggc tac cag acc agc aaa gcg Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala 210 215 220	672

aac tcg tat caa aaa ccg aca gag aag cag tac ctc tag
 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

711

5
 <210> 10
 <211> 237
 <212> PRT
 <213> Künstliche Sequenz
 10 <223> Beschreibung der künstlichen Sequenz:BB5

<400> 10
 Met Arg Ile Leu Val Ile Asn Pro Asn Ser Ser Ser Ala Leu Thr Glu
 1 5 10 15

15 Ser Val Ala Asp Ala Ala Gln Gln Val Val Ala Thr Gly Thr Ile Ile
 20 25 30

20 Ser Ala Ile Asn Pro Ser Arg Gly Pro Ala Val Ile Glu Gly Ser Phe
 35 40 45

Asp Glu Ala Leu Ala Thr Phe His Leu Ile Glu Glu Val Glu Arg Ala
 50 55 60

25 Glu Arg Glu Asn Pro Pro Asp Ala Tyr Val Ile Ala Cys Phe Leu Asp
 65 70 75 80

Pro Gly Leu Asp Ala Val Lys Glu Leu Thr Asp Arg Pro Val Val Gly
 85 90 95

30 Val Ala Glu Ala Ala Ile His Met Ser Ser Phe Val Ala Ala Thr Phe
 100 105 110

35 Ser Ile Val Ser Ile Leu Pro Arg Val Arg Lys His Leu His Glu Leu
 115 120 125

Val Arg Gln Ala Gly Ala Thr Asn Arg Leu Ala Ser Ile Lys Leu Pro
 130 135 140

Asn Leu Gly Val Met Ala Phe His Glu Asp Glu His Ala Ala Leu Glu
 145 150 155 160

45 Thr Leu Lys Gln Ala Ala Lys Glu Ala Val Gln Glu Asp Gly Ala Glu
 165 170 175

Ser Ile Val Leu Gly Cys Ala Gly Met Val Gly Phe Ala Arg Gln Leu
 180 185 190

50 Ser Asp Glu Leu Gly Val Pro Val Ile Asp Pro Val Glu Ala Ala Cys
 195 200 205

Arg Val Ala Glu Ser Leu Val Ala Leu Gly Tyr Gln Thr Ser Lys Ala
 210 215 220

55 Asn Ser Tyr Gln Lys Pro Thr Glu Lys Gln Tyr Leu
 225 230 235

<211> 25
<212> DNA
<213> Künstliche Sequenz

5 <220>
<223> Beschreibung der künstlichen Sequenz: Primer5

<400> 11
gccccaaaggaa atgggtgcattt catcg 25

10

<210> 12
<211> 30
<212> DNA
15 <213> Künstliche Sequenz

<220>
<223> Beschreibung der künstlichen Sequenz: Primer6

0 <400> 12
ggtcagggtgg gttccaccgcg ctactgccgc 30

25

<210> 13
<211> 5777
<212> DNA
<213> Künstliche Sequenz

30 <220>
<223> Beschreibung der künstlichen Sequenz: Plasmid pOM21

<400> 13
aattcttaag aaggagatatacatatgaga atcctcgta tcaaccccaa cagttccagc 60

35 gccccttactg aatcggttgc ggacgcagca caacaagggttgc tcgcgaccgg caccataatt 120
tctgccatca accccctccag aggacccgcg gtcattgaag gcagctttga cgaagcactg 180
gccacgttcc atctcattga agaggtggag cgcgctgagc gggaaaaccc gcccgcacgc 240
tacgcatcg catgtttcggttggatccggc cttgacgcgg tcaaggagct gactgacagg 300
ccagtggttag gagttgccga agctgcaatc cacatgtttt cattcgatcg ggcacaccc 360

45 tccattgtca gcatcctccc gagggtcagg aaacatctgc acgaactggc acggcaagcg 420
ggggcgacga atcgccctcgcc ctccatcaag ctcccaatc tgggggtgat ggccttccat 480
50 gaggacgaac atgcccgcact ggagacgctc aaacaaggccg ccaaggaggc ggtccaggag 540
gacggcgccg agtcgatagt gtcggatgc gcccggatgg tgggggttgc gctcaactg 600
agcgacgaac tcggcgatccc tgcattcgac cccgtcgagg cagttgcgg cgtggccgag 660

55 agtttggtcg ctctggctta ccagaccaggc aaagcgaact cgtatcaaaa accgacacag 720
aagcagttacc tctagctgca gccaagcttc tgggggttgc gatgagagaa gatggatcg 780
ctgatacaga ttaaatcaga acgcagaaggc ggtctgataa aacagaattt ggcctggcg 840

agtagcgcgg tggtcccacc tgacccatg ccgaactcag aagtgaaacg ccgtagcgcc 900
5 gatggtagtg tggggctc ccatgcgaga gtaggaaact gccaggcatc aaataaaacg 960
aaaggctcag tcgaaagact gggccttcg ttttatctgt tgtttgcgg tgaacgctct 1020
cctgagtagg acaaattccgc cgggagcggg tttgaacgtt gcgaagcaac ggccggagg 1080
10 gtggcgggca ggacgcccgc cataaaactgc caggcatcaa attaaggcaga aggccatcct 1140
gacggatggc cttttgcgt ttctacaaac tctttgttt attttctaa atacattcaa 1200
atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat cgtccattcc 1260
15 gacagcatcg ccagtcacta tggcgtgctg cttagcgctat atgcgttgat gcaattcta 1320
tgcgcacccg ttctcgagc actgtccgac cgcttggcc gcccggagt cctgctcgct 1380
0 tcgctacttg gagccactat cgactacgcg atcatggcga ccacacccgt cctgtggatc 1440
ctctacgccc gacgcacgt ggcggcactc accggcgcca caggtgcggg tgctggcgcc 1500
tatatcgccg acatcaccga tgggaaagat cgggctcgcc acttcgggct catgagcgct 1560
25 tgtttggcg tgggtatggt ggcaggcccc gtggccgggg gactgttggg cgccatctcc 1620
ttgcatgcac cattccttgc ggcggcggtg ctcaacggcc tcaacctact actggctgc 1680
30 ttccataatgc aggagtgcga taaggagag cgtcgaccga tgccctttag agccttcaac 1740
ccagtcagct cttccgggtg ggcgcggggc atgactatcg tcgcccact tatgactgtc 1800
ttcttatca tgcaactcgt aggacaggtg cggcagcgc tctgggtcat ttcggcgag 1860
35 gaccgcttc gctggagcgc gacgatgatc ggcctgtcgc ttgcgttatt cggaatcttgc 1920
cacgcctcg ctcaagcctt cgtcactggt cccgcccacca aacgtttcggt cgagaagcag 1980
gccattatcg ccggcatggc ggccgacgcg ctgggctacg tcttgctggc gttcgcgacg 2040
cgaggctgga tggcctccc cattatgatt cttctcgctt cggcggcat cgggatgccc 2100
45 gcgttgcagg ccatgctgtc caggcaggta gatgacgacc atcagggaca gcttcaagga 2160
tcgctcgccg ctcttaccag cctaacttcg atcactggac cgctgatcgt cacggcgatt 2220
tatgccgcct cggcgagcac atggaacggg ttggcatgga ttgtaggcgc cgccctatac 2280
50 cttgtctgcc tccccgcgtt gcgtcgccgt gcatggagcc gggccacctc gacctgaatg 2340
gaagccggcg gcacctcgct aacggattca ccactccaag aattggagcc aatcaattct 2400
tgccggagaac tgtaatgcg caaaccacc cttggcagaa catatccatc gctccgcac 2460
55 tctccagcag ccgcacgcgg cgcatctcgg gcagcggttgg gtcctggcca cgggtgcgc 2520
tgatcgtgct cctgtcggttgg aggacccggc taggctggcg gggttgcctt actggtagc 2580

agaatgaatc accgatacgc gagcgaacgt gaagcgactg ctgctgaaa acgtctgcga 2640
 cctgagcaac aacatgaatg gtcttcggtt tccgtgttc gtaaagtctg gaaacgcgga 2700
 5 agtccccctac gtgctgctga agttgcccgc aacagagagt ggaaccaacc ggtgatcacca 2760
 cgatactatg actgagagtc aacgccatga gcggcctcat ttcttattct gagttacaac 2820
 10 agtccgcacc gctgtccggt agtccttcc ggtgggcgcg gggcatgact atcgtcgccg 2880
 cacttatgac tgcttcctt atcatgcaac tcgttaggaca ggtgccggca ggcaccaaca 2940
 gtccccccggc cacggggcct gccaccatac ccacgcccga acaagcgccc tgcaccatta 3000
 15 tggccggat ctgcatcgca ggatgctgct ggctaccctg tggaaacacct acatctgtat 3060
 taacgaagcg ctaaccgttt ttatcaggct ctgggaggca gaataaatga tcataatcg 3120
 0 aattattacc tccacgggga gaggctgagc aaactggcct caggcattt agaagcacac 3180
 ggtcacactg cttccggtag tcaataaacc ggtaaaccag caatagacat aagcggctat 3240
 ttaacgaccc tgccctgaac cgacgaccgg gtcgaattt ctttgcatt tctgcattc 3300
 25 atccgcttat tatcacttat tcaggcgtag caccaggcgt ttaagggcac caataactgc 3360
 cttaaaaaaa ttacgccccg ccctgccact catcgca gtaatgtat tcattaagca 3420
 ttctgcccac atgaaagcca tcacagacgg catgtgaac ctgaatcgcc agcggcatca 3480
 30 gcaccttgac gccttgcgtta taatatttgc ccatggtaaa aacggggcgc aagaagttgt 3540
 ccatattggc cacgtttaaa tcaaaaactgg tgaaaactcac ccagggattt gctgagacga 3600
 35 aaaaacatatt ctcaataaac cctttaggaa aataggccag gtttcaccg taacacgcca 3660
 catcttgcga atatatgtgt agaaaactgac ggaaatcg 3720
 atgaaaacgt ttcaatggc tcatggaaaa cgggttaaca agggtaaca ctatccata 3780
 tcaccagctc accgtcttcc attgcatac gaattccgga tgagcattca tcaggcggc 3840
 aagaatgtga ataaaggccg gataaaactt gtgcttattt ttcttacgg tctttaaaaa 3900
 45 ggccgtaata tccagctgaa cggctgggtt ataggtacat tgagcaactg actgaaatgc 3960
 ctcaaaatgt tcttacgat gcccattggaa tataatcaacg gtggtatatc cagtgatttt 4020
 50 tttctccatt ttagcttcct tagctcctga aaatctcgat aactcaaaaa atacgcccgg 4080
 tagtgatctt atttcattat ggtgaaagtt ggaacctt acgtgcccgt caacgtctca 4140
 ttttcgccaa aagttggccc agggcttccc ggtatcaaca gggacaccag gatttattta 4200
 55 ttctgcgaag tgatcttcgg tcacaggat ttatcggcg caaagtgcgt cgggtgatgc 4260
 tgccaaactta ctgatttatgt gtatgatggt gtttttgagg tgctccagtg gcttctgttt 4320
 ctatcagctg tccctcctgt tcagctactg acggggtggt gcgttaacggc aaaagcaccg 4380

ccggacatca ggcgtacgg agtgtatact ggcttactat gttggcactg atgaggggt 4440
 5 cagtgaaatg cttcatgtgg caggaaaaa aaggctgcac cggtcgtca gcagaatatg 4500
 tgatacagga tatattccgc ttccctcgctc actgactcgc tacgctcggt cgttcgactg 4560
 cggcgagcgg aaatggctta cgaacggggc ggagattcc tggaagatgc caggaagata 4620
 10 cttaaacaggg aagtgagagg gcccggcaa agccgtttt ccataggctc cgccccctg 4680
 acaaggcatca cgaatctga cgctcaaatc agtggtggcg aaacccgaca ggactataaa 4740
 15 gataccaggc gttccctg gcccctc cgtgcgtct cctgtccctg ctttcgggt 4800
 taccgggtgc attccgtgt tatggccgcg tttgtctcat tccacgcctg acactcagtt 4860
 0 cccggtaggc agtgcgtcc aagctggact gtatgcacga accccccgtt cagtccgacc 4920
 gtcgcgcctt atccgtaac ttcgtcttg agtccaaccc ggaaagacat gcaaaagcac 4980
 cactggcage agccactggt aattgattt gaggagttt tcttgaagtc atgcgcgg 5040
 25 taaggctaaa ctgaaaggac aagttttggt gactgcgtc ctccaagcca gttacctcg 5100
 ttcaaagagt tggtagctca gagaaccttc gaaaaaccgc cctgcaaggc ggtttttcg 5160
 ttttcagagc aagagattac ggcagacca aaacgatctc aagaagatca tcttattaaat 5220
 30 cagataaaat attcaagat ttcagtgc aaatcttcaa caaatgtac acctgaagtc 5280
 agccccatac gatataagtt gtaattctca ttttgacag cttatcatcg ataagttta 5340
 35 atgcggtagt ttatcacagt taaattgcta acgcgtc acgcgtgtaa tgaaatctaa 5400
 caatgcgtc atcgtcatcc tggcaccgt caccctggat gctgtaggca taggcttgg 5460
 tatgccggta ctgccggcc tcttgcggga ttagtcatgc cccgcgcacca ccggaaaggag 5520
 ctgactgggt tgaaggctct caagggcatc ggtcgacgt ctcccttatg cgactcctgc 5580
 attaggaagc agcccgatg taggttgagg ccgttgagca ccggccgc aaggaatgg 5640
 gcatgcacatcg atcaccacaa ttcagcaat ttttgcacatc atcagttca ttttcctg 5700
 45 gttgcacatg gcccatttc ctgtcagtaa cgagaaggc gcaatttcg ggcctttta 5760
 gactggtcgt aatgaac 5777
 50 <210> 14
 <211> 7175
 <212> DNA
 <213> Künstliche Sequenz
 55 <220>
 <223> Beschreibung der künstlichen Sequenz: Plasmid pOM22
 <400> 14

aattcttaag aaggagatat acatatgacc ctgcagaaaag cgcaagcgna gcgcatttag 60
aaagagatct gggagcttc cccggttctcg gcggaaggcc cccgtgttac cccgctgacc 120
5 tacactccag agcatgccgc cgccggaa acgctcattt cggctatgga agcggccgct 180
ttgagcgttc gtgaagacgc tctcggaaac atcatcgcc gacgtgaagg cactgatccg 240
10 cagctccctg cgatcgccgt cggttacac ttcgattctg tccgaaacgg cgggatgttc 300
gatggcactg caggcgtggt gtgcgcctt gaggctgccc gggtgatgct ggagagcggc 360
tacgtgaatc ggcattccatt ttagttcatc gcgatcgtgg aggaggaagg ggcccgcttc 420
15 agcagtggca tggtggcgg ccggccatt gcagggttgg tcgccgacag ggaactggac 480
tctttgggtt atgaggatgg agtgtccgtt aggcaggcgg ctactgcctt cggcttgaag 540
0 ccgggcgaac tgcaggctgc agccgcctcc gcggcggacc tgcgtgctt tatcgaacta 600
cacattgaac aaggaccgat cctcgagcag gagcaaataag agatcgaggat tgtgaccctcc 660
atcggtggcg ttgcgcatt gcgggttgct gtcaaaggca gaagcgcaca cggccggcaca 720
25 acccccatgc acctgcgcaca ggatgcgctg gtacccgcgg ctctcatggc gcgggagggtc 780
aaccggttcg tcaacgagat cggcgatggc acagtggcta ccgttggcca cctcacagt 840
gcccccggtg gcggcaacca ggtccgggg gaggtggagt tcacactgga cctgcgttct 900
30 cgcgtgagg agtcgctccg ggtgttgatc aaccgcattt cggtcattt cggcgagggtc 960
gcctcgagg ccgggtgtggc tgccgatgtt gatgaatttt tcaatctcag cccgggtgcag 1020
35 ctggctccata ccatggtgga cggcggttcgc gaagcggcct cggccctgca gttcacgcac 1080
cgggatataca gcagtggggc gggccacgc tcgatgttca tcgcccaggat cacggacgtc 1140
ggaatggttt tcgttccaag ccgtgctggc cggagccacg ttcccaaga atggaccgat 1200
ttcgatgacc ttgcgaaggaa aactgagggtt gtcctccggg taatgaaggc acttgaccgg 1260
ggatcccatc atcatcatca tcattgactg cagccaaatct tctgtttgg cggatgagag 1320
45 aagatttca gcctgataca gattaaatca gaacgcagaa gcggtctgtat aaaacagaat 1380
ttgcctggcg gcagtagcgc ggtggtccca cctgacccca tgccgaactc agaagtggaa 1440
cgccgtagcg ccgatggtag tgtgggtct cccatgcga gagtagggaa ctgccaggca 1500
50 tcaaataaaa cgaaaggctc agtcgaaaga ctgggcctt cgttttatct gttgtttgtc 1560
ggtgaacgct ctccctgagta ggacaaatcc gcccggagcg gatttgaacg ttgcgaagca 1620
55 acggcccgga ggggtggcggg caggacgccc gccataaact gccaggcatc aaattaagca 1680
gaaggccatc ctgacggatg gccttttgc gtttctacaa actctttgt ttatcttct 1740
aaatacattc aaatatgtat ccgctcatga gacaataacc ctgataaaatg cttcaataat 1800

attaaaaaag gaagagtatg agtattcaac attccgtgt cgcccttatt cccttttg 1860
5 cggcattttg cttcctgtt tttgctcacc cagaaacgct ggtgaaagta aaagatgctg 1920
aagatcagtt ggggcacga gtgggttaca tcgaactgga tctcaacagc ggtaagatcc 1980
ttgagagtt tcgccccgaa gaacgtttc caatgatgag cactttaaa gttctgctat 2040
10 gtggcgcggt attatccgt gttgacgccc ggcaagagca actcggtcgc cgcatacact 2100
attctcagaa tgacttgggtt gagtactcac cagtcacaga aaagcatctt acggatggca 2160
15 tgacagtaag agaattatgc agtgctgcca taaccatgag tgataacact gcccggaaact 2220
tacttctgac aacgatcggaa ggaccgaagg agctaaccgc tttttgcac aacatgggg 2280
atcatgtAAC tcgccttgat cgttggaaac cggagctgaa tgaagccata ccaaacgacg 2340
0 agcgtgacac cacgatgcct gtagcaatgg caacaacgtt gcgcaaaacta ttaactggcg 2400
aactacttac tctagcttcc cggcaacaat taatagactg gatggaggcg gataaagttg 2460
25 caggaccact tctgcgctcg gccctccgg ctggctggtt tattgctgat aaatctggag 2520
ccggtgagcg tgggtctcgc ggtatcattg cagcactggg gccagatggt aagccctccc 2580
gtatcgtagt tatctacacg acggggagtc aggcaactat ggatgaacga aatagacaga 2640
30 tcgctgagat agtgccctca ctgattaagc attggtaact gtcagaccaa gtttactcat 2700
atatacttta gattgattta aaacttcatt tttaatttaa aaggatctag gtgaagatcc 2760
35 ttttgataa tctcatgacc aaaatccctt aacgtgagtt ttcgttccac tgagcgtcag 2820
accccgtaga aaagatcaaa ggatcttctt gagatccttt ttttctgcgc gtaatctgct 2880
gcttgcaaac aaaaaaaacca cggctaccag cgggggtttg tttgccggat caagagctac 2940
caactcttt tccgaaggta actggcttca gcagagcgc gataccaaat actgtccctc 3000
tagttagcc gtagttaggc caccattca agaactctgt agcaccgcct acataccctcg 3060
45 ctctgctaattt cctgttacca gtggctgctg ccagtggcga taagtcgtgt cttaccgggt 3120
tggactcaag acgatagttt ccggataagg cgcagcggtc gggctgaacg ggggggttcgt 3180
gcacacagcc cagcttggag cgaacgacct acaccgaact gagataccta cagcgtgagc 3240
50 tatgagaaag cgccacgctt cccgaaggga gaaaggcggga caggtatccg gtaagcggca 3300
gggtcggaac aggagagcgc acgagggagc ttccaggggg aaacgcctgg tatcttata 3360
55 gtcctgtcgg gtttcgccac ctctgacttg agcgtcgatt tttgtatgc tcgtcagggg 3420
ggcggagcct atggaaaaac gccagcaacg cggcctttt acggttcctg gcctttgct 3480
ggcctttgc tcacatgttc ttcctgcgt tatccctga ttctgtggat aaccgtattta 3540

ccgccttga gtgagctgat accgctcgcc gcagccgaac gaccgagcgc agcgagtcag 3600
tgagcgagga agcggaaagag cgccctgatgc ggtattttct ccttacgcat ctgtcggtta 3660
5 ttccacacccg catatatggt gcactctcag tacaatctgc tctgatgccg catagttaaag 3720
ccagtataca ctccgctatc gctacgtgac tgggtcatgg ctgcgccccg acacccgcca 3780
10 acacccgctg acgcccctg acgggcttgc ctgctcccg catccgctta cagacaagct 3840
gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgtcatcacc gaaacgcgcg 3900
aggcagctgc ggtaaagctc atcagcgtgg tctgtgaagcg attcacagat gtctgcctgt 3960
15 tcatccgcgt ccagctcggtt gagtttctcc agaagcgtta atgtctggct tctgataaag 4020
cgggccccatgt taagggcggt ttttctgt ttggtcactt gatgcctccg tgtaaggggg 4080
0 aatttctgtt catgggggtta atgataccga tgaaacgaga gaggatgctc acgataacggg 4140
ttactgtatga tgaacatgcc cggttactgg aacgttgtga gggtaaacaa ctggcggtat 4200
ggatgcggcg ggaccagaga aaaatcactc agggtcaatg ccagcgcctc gttaatacag 4260
25 atgttaggtgt tccacagggt agccagcagc atcctgcgtat gcagatccgg aacataatgg 4320
tgcagggcgcc tgacttccgc gttccagac tttacgaaac acggaaaccg aagaccatcc 4380
atgttggtgc tcaggtcgca gacgtttgc agcagcagtc gcttcacgtt cgctcggtta 4440
30 tcgggtgattc attctgctaa ccagtaaggc aacccgcca gcctagccgg gtccctcaacg 4500
acaggagcac gatcatgcgc acccgtggcc aggacccaac gctgcccggatgcgcgcg 4560
35 tgccggctgct ggagatggcg gacgcgtatgg atatgttctg ccaagggttg gtttgcgtat 4620
tcacagttct ccgcaagaat tgattggctc caattcttgg agtggtaat ccgttagcga 4680
ggtgcggcccg gcttccatttc aggtcgaggt ggccccggctc catgcaccgc gacgcaacgc 4740
ggggaggcag acaaggatata gggcggcgcc tacaatccat gccaacccgt tccatgtgct 4800
cgccgaggcg gcataaaatcg ccgtgacgtat cagcggtcca gtgatcgaag ttaggcttgt 4860
45 aagagccgcg agcgatcctt gaagctgtcc ctgatggctg tcatctacat gcctggacag 4920
catggcctgc aacgcgggca tcccgatgcc gccggaaagcg agaagaatca taatggggaa 4980
ggccatccag cctcgctcg cgaacgcac caagacgtat cccagcgcgt cggccgcatt 5040
50 gcccggcata atggcctgtt tctcgccaa acgtttgggt gcgggaccag tgacgaaggc 5100
ttgagcgagg gcgtgcaaga ttccgaatac cgcaagcgcac agggcgtatca tcgtcggtat 5160
55 ccagcggaaag cggtcctcg cggaaatgc acagagcgtt gcccggacat gtccctacgag 5220
ttgcgtatgata aagaagacag tcataagtgc ggcgacgtat gtcgtatggcc ggcggccacccg 5280
gaaggagctg actgggttga aggctctaa gggcatcggt cgacgcgttc ccttatgcga 5340

5' ctcctgcatt aggaagcagc ccagtagtag gttgaggccg ttgagcaccg ccggccgcaag 5400
 5 gaatggtgca tgcacatcgatc accacaattc agcaaattgt gaacatcatc acgttcatct 5460
 ttcctgggtt gccaatggcc cattttcctg tcagtaacga gaaggtcgcg aattcaggcg 5520
 ctttttagac tggtcgtaat gaacaattct taagaaggag atatacatat gtttgacgta 5580
 10 atagtttaaga actgcccgtat ggtgtccagc gacggaatca ccgaggcaga cattctggtg 5640
 aaagacggca aagtgcgcgc aatcagctcg gacacaagtg atgttgaggc gagccgaacc 5700
 15 attgacgcgg gtggcaagtt cgtgatgccc ggcgtggcg atgaacatgt gcatatcatc 5760
 gacatggatc tgaagaacccg gtatggccgc ttcgaactcg attccgagtc tgcggccgtg 5820
 ggaggcatca ccaccatctt tgagatgccc tttaccttcc cgcccaccac cactttggac 5880
 20 gccttcctcg aaaagaagaa gcaggcgggg cagcggttga aagttgactt cgcgctctat 5940
 ggccggtggag tgccggaaa cctgcccagc atccgcaaaa tgcacgacgc cggcgcagtg 6000
 25 ggcttcaagt caatgatggc agcctcagtt ccgggcattgt tcgacgcccgt cagcgacggc 6060
 gaactgttcg aaatcttcca ggagatcgca gcctgtggtt cagtcgcccgt ggtccatgcc 6120
 gagaatgaaa cgatcattca agcgctccag aagcagatca aagccgctgg tcgcaaggac 6180
 30 atggccgcct acgaggcatc ccaaccagtt ttccaggaga acgaggccat tcagcgtcg 6240
 ttactactgc agaaagaagc cggctgtcga ctgattgtgc ttcacgtgag caaccctgac 6300
 35 ggggtcgagc tgatacatcg ggccaaatcc gagggccagg acgtccactg cgagtcgggt 6360
 ccgcagtatc tgaatatcac cacggacgac gccgaacgaa tcggaccgtat tatgaaggtc 6420
 gcccggcccg tccgctcagc cgagatgaac gtcagattat gggaaacaact tgagaacggg 6480
 ctcatcgaca cccttgggtc agaccacggc ggacatcctg tcgaggacaa agaaccggc 6540
 tggaaggacg tggaaagc cggcaacggc ggcgtggcc ttgagacatc cctgcctatg 6600
 45 atgctgacca acggagtgaa taaaggcagg ctatccttgg aacgcctcgt cgaggtgatg 6660
 tgcgagaaac ctgcgaagct ctttggcatc tatccgcaga agggcacgct acaggttggt 6720
 tccgacgccc atctgctcat cctcgatctg gatattgaca ccaaagtggc tgcctcgcag 6780
 50 ttccgatccc tgcataagta cagccgttc gacggatgc ccgtcacggg tgcaccgggt 6840
 ctgacgatgg tgcgcggaac ggtggggca gagaagggag aagttctggc cgagcaggga 6900
 55 ttcggccagt tcgtcaccgg tcacgactac gaggcgatcg agtgaggatc tcgacgctct 6960
 cccttatgcg actcctgcatt taggaagcag cccagtagta ggttgaggcc gttgagcacc 7020
 gccggccgca ggaatggtgc atgcacatcgat caccacaatt cagcaaattt gtaacatcat 7080

cacgttcatc ttccctggc tgccaatggc ccattttcct gtcagtaacg agaaggcgc 7140
gaattcaaggc actttttaga ctggtcgtaa tgaac 7175

5
<210> 15
<211> 5989
<212> DNA
<213> Künstliche Sequenz

10 <220>
<223> Beschreibung der künstlichen Sequenz:Plasmid pDHYH

15 <400> 15
aattcttaag aaggagatat acatatggat gcaaagctac tggttggcgg caactattttt 60
tcctcgaccc gcaaaaatccg agccgacgtg ctgattgaaa acggcaaagt cggccgtgtc 120
ggcatgctgg acgccccgac gccggacaca gttgagcggg ttgactgcga cggcaaatac 180
gtcatgccccg gcggtatcga cgttcacacc cacatcgact cccccctcat ggggaccacc 240
accggccqata atttttgtcaq cgaaacgatt gcagccgcta ccggcggaaac aacgaccatc 300

25 gtcgatttcg gacagcagct cgccggcaag aacctgctgg aatccgcaga cgcgaccac 360
aaaaaggcgc agggaaatc cgtcattgat tacggcttcc atatgtgcgt gacgaacctc 420
..... 480

30 gtcttcatgg cctaccgcgg aagcctgatg atcaacgacg gcgaactgtt cgacatcctc 540
aaggagatcg gctccacgggg tgccaaacta tccgtccacg cagagaacgg cgacgtcatc 600

35 gacaggatcg ccgcccacct ctacgccccaa ggaaaaaaccg ggcccccggac ccacgagatc 660
gcacgccccgc cggaatcgga agtcgaagca gtcagccggg ccatcaagat ctcccgatg 720

gccgaggtgc cgctgtattt cgtgcattt tccacccagg gggccgtcga ggaagtagct 760
gccgcgcaga tgacaggatg gccaatcagc gccgaaacgt gcaccacta cctgtcgctg 840
agccgggacaca tctacgacca gccgggattc gagccggcca aagctgtcct cacaccacccg 900

45 ctgcgcacac aggaacacca ggacgcgttg tggagaggca ttaacaccgg tgcgctcagc 960
gtcgtcagtt ccgaccactg ccccttctgc tttgaggaaa agcagcggat gggggcagat 1020

50 gacttccggc agatccccaa cggcgggccc ggcgtggagc accgaatgtc cgtgtatgtat 100
gagaccggtg tcgcggaagg aaaaatgacg atcgagaaaat tcgtcgaggt gactgccgag 1140
aacccgggca agcaattcga tatgtacccg aaaaagggaa caattgcacc gggctccgat 1200

55 gcagacatca tcgtggtcga ccccaacgga acaaccctca tcagtgccga cacccaa 1260
caaaacatgg actacacgct gttcgaaggc ttcaaaatcc gttgctccat cgaccaggtg 1320
ttctcgcggtg gcgacctgat cagcgtcaaa ggcgaatatg tcggcacccg cggccgcggc 1380

gaattcatca agcgaggcgc ttggagccac ccgcagttcg aaaaataaaa gcttggctgt 1440
5 tttggcggat gagagaagat ttcagcctg atacagatta aatcagaacg cagaagcggt 1500
ctgataaaac agaatttgc tggcggcagt agcgcggtgg tcccacctga ccccatgccc 1560
aactcagaag tgaaaacgccc tagcgccgat ggttagtgtgg ggtctccca tgcgagagta 1620
10 gggactgcc aggcataaa taaaacgaaa ggctcagtcg aaagactggg ccttcgttt 1680
tatctgttgt ttgtcggtga acgctctcct gagtaggaca aatccgcccgg gagcggattt 1740
15 gaacgttgcg aagcaacggc ccggagggtg gggggcagga cgcccgccat aaactgcccag 1800
gcatcaaatt aagcagaagg ccatcctgac ggatggcctt tttgcgttac tacaaactct 1860
tttgtttatt ttctaaata cattcaaata tgcgttccgtt catgagacaa taaccctgtat 1920
0 aaatgcttca ataataattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc 1980
ttattccctt tttgcggca tttgccttc ctgttttgc tcacccagaa acgctggtga 2040
25 aagtaaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa ctggatctca 2100
acagcggtaa gatccttgag agtttcgccc ccgaagaacg tttccaatg atgagcactt 2160
ttaaagttct gctatgtggc gcggtattat cccgtgttga cgccgggcaa gagcaactcg 2220
30 gtcggccat acactattct cagaatgact tggttgagta ctcaccagtc acagaaaagc 2280
atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata 2340
35 acactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta accgctttt 2400
tgcacaacat gggggatcat gtaactcgcc ttgatcggtt ggaaccggag ctgaatgaag 2460
ccataccaaa cgacgagcgt gacaccacga tgcctgttagc aatggcaaca acgttgccca 2520
aactattaac tggcgaacta cttactctag cttccggca acaattaata gactggatgg 2580
45 aggccgataa agttgcagga ccacttctgc gctcgccct tccggctggc tggtttattt 2640
ctgataaaatc tggagccggt gagcgtgggt ctcgcgttat cattgcagca ctggggccag 2700
atggtaagcc ctcccgatc gtagttatct acacgacggg gagtcaggca actatggatg 2760
aacgaaatag acagatcgat gagataggtg cctcactgtat taagcattgg taactgtcag 2820
50 accaagttt ctcataatata ctttagattt atttaaaact tcattttaa tttaaaagga 2880
tcttaggtgaa gatcctttt gataatctca tgacaaaaat cccttaacgt gagtttcgt 2940
55 tccactgagc gtcagacccc gtagaaaaaga tcaaaggatc ttctttagat ctttttttc 3000
tgcgcgtaat ctgctgcttgc caaacaaaaa aaccaccgct accagcgggtg gtttgggttgc 3060
cgatcaaga gctaccaact cttttccga agttaactgg cttcagcaga ggcgcagatac 3120

caaatactgt cttcttagtg tagccgtagt taggccacca cttcaagaac tctgttagcac 3180
cgccata cctcgctctg ctaatcctgt taccagtggc tgctgccagt ggcgataagt 3240
5 cgtgtcttac cgggttggac tcaagacgt agttaccgga taaggcgcag cggtcggct 3300
gaacgggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc gaactgagat 3360
10 acctacagcg tgagctatga gaaagcgcca cgcttcccgaa agggagaaag gcggacaggt 3420
atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg 3480
cctggtatct ttatagtcct gtcgggttcc gccacctctg acttgagcgt cgattttgt 3540
15 gatgctcgtc aggggggcgg agcctatgga aaaacgcccag caacgcggcc ttttacggt 3600
tcctggcctt ttgctggcct tttgctcaca tgttcttcc tgcgttatcc cctgattctg 3660
20 tggataaccg tattaccgccc tttgagtgag ctgataaccgc tcgccgcagc cgaacgcaccg 3720
agcgcagcga gtcagtgagc gaggaagcgg aagagcgcct gatgcggtat tttctcctta 3780
cgcatctgtg cggttattca caccgcatac atggtgcact ctcagtcaca tctgctctga 3840
25 tgccgcatacg ttaagccagt atacactccg ctatcgctac gtgactgggt catggctgcg 3900
ccccgacacc cgccaacacc cgctgacgcg ccctgacggg cttgtctgct cccggcatcc 3960
gcttacagac aagctgtgac cgtctccggg agctgcattt gtcagagggtt ttcaccgtca 4020
30 tcaccgaaac gcgcgaggca gctgcggtaa agctcatcag cgtggctgt aagcgattca 4080
cagatgtctg cctgttcattc cgctccagc tcgttgagtt tctccagaag cgtaatgtc 4140
35 tggcttctga taaagcgggc catgttaagg gcggttttt cctgtttgggt cacttgatgc 4200
ctccgtgtaa gggggattt ctgttcatgg gggtaatgtat accgatgaaa cgagagagga 4260
tgctcacgt acgggttact gatgatgaac atgcccgggtt actggaacgt tgtgagggt 4320
aacaactggc ggtatggatg cggcgggacc agagaaaaat cactcagggt caatgcgcagc 4380
gcttcgttaa tacagatgta ggtgttccac agggtagcca gcagcatcct gcgtatgcaga 4440
45 tccggaacat aatggtgcag ggccgtgact tccgcgttcc cagactttac gaaacacgga 4500
aaccgaagac cattcatgtt gttgctcagg tcgcagacgt tttgcagcag cagtcgcttc 4560
acgttcgctc gcgtatcggt gattcattct gctaaccagt aaggcaaccc cgccagccata 4620
50 gcccgggtcct caacgcacagg agcacgatca tgccgcacccg tggccaggac ccaacgcgtc 4680
ccgagatgcg ccgcgtgcgg ctgctggaga tggcggacgc gatggatatg ttctgccaag 4740
55 ggttggtttgcgcattcaca gtttccgcata agaattgatt ggctccaatt cttggagtgg 4800
tgaatccgtt agcgaggtgc cgccggcttc cattcagggtc gaggtggccc ggctccatgc 4860
accgcgcacgc aacgcgggaa ggcagacaag gtataggcgc gcgcctacaa tccatgccaa 4920

cccgttccat gtgctcgccg aggcggcata aatcgccgtg acgatcagcg gtccagtgt 4980
 5 cgaagttagg ctggtaagag ccgcgagcga tccttgaagc tgtccctgat ggtcgcatc 5040
 tacctgcctg gacagcatgg cctgcaacgc gggcatcccg atgcccggg aagcgagaag 5100
 aatcataatg gggaaaggcca tccagcctcg cgtcgcgaac gccagcaaga cgtagcccag 5160
 10 cgcgtcgcc gccatgccgg cgataatggc ctgcttctcg ccgaaacgtt tggtggcggg 5220
 accagtgacg aaggctttag cgagggcgtg caagattccg aataccgcaa gcgcacaggcc 5280
 15 gatcatcgtc gcgcgtccagc gaaagcggtc ctcgcccggg atgaccaga ggcgtgcccgg 5340
 cacctgtcct acgagttgca tgataaagaa gacagtcata agtgcggcga cgatagtcat 5400
 gccccgcgcc caccggaaagg agctgactgg gttgaaggct ctcaaggcga tcggtcgacg 5460
 20 ctctccctta tgcgactcct gcattaggaa gcagcccagt agtaggttga ggccgttgag 5520
 caccggccgc gcaaggaatg gtgcattgtc gatggctacg agggcagaca gtaagtggat 5580
 ttaccataat cccttaattt tacgcaccgc taaaacgcgt tcagcgcat cacggcagca 5640
 25 gacaggtaaa aatggcaaca aaccacccta aaaactgcgc gatcgccct gataaatttt 5700
 aaccgtatga atacctatgc aaccagaggg tacaggccac attacccca cttaatccac 5760
 30 tgaagctgcc atttttcatg gtttaccat cccagcgaag ggccatgcgt gcatcgaaat 5820
 taatacgacg aaattaatac gactcactat agggcaattt cgatcaccac aattcagcaa 5880
 attgtgaaca tcatcacgtt catcttccc tggttgccaa tggcccttt tcctgtcagt 5940
 35 aacgagaagg tcgcgaattt aggcgtttt tagactggc gtaatgaac 5989

<210> 16
 <211> 6958
 <212> DNA
 <213> Künstliche Sequenz

45 <220>
 <223> Beschreibung der künstlichen Sequenz: Plasmid
 pJAVIER16

<400> 16
 aattcttaag aaggagatat acatatggcg aaaaacttga tgctcgccgt cgctcaagtc 60
 50 ggcggtatcg atagttcgga atcaagaccc gaagtcgtcg cccgcttgc tggccctgctg 120
 gaagaagcag cttcccaggg cgccgaactg gtggtcttc ccgaactcac gctgaccacg 180
 55 ttcttccgc gtacctgggt cgaagaaggc gacttcgagg aataacttgc taaatccatg 240
 cccaatgacg acgtcgccgc cttttcgaa cgcccaaag accttggcgt gggcttctac 300
 ctcggatacg cgaaactgac cagtgatgag aagcggtaca acacatcaat tctggtaac 360

aagcacggcg acatcgtcgg caagtaccgc aagatgcac tgccgggcca cgccgataac 420
5 cgggaaggac tacccaacca gcaccttgaa aagaaaatact tccgcgaagg agatctcgga 480
ttcggtgtct tcgacttcca cggcgtgcag gtcggaatgt gtatctgcaa cgaccggcga 540
tggccggagg tctaccgctc tttggccctg cagggagcag agctcgctgt cctgggctac 600
10 aacaccccg atttcgttcc cggctggcag gaagagcctc acgcaagat gttcacgcac 660
cttcttcac ttcaggcagg ggcataccag aactcggtat ttgtggcggc tgccggcaag 720
15 tcgggcttcg aagacgggc a ccacatgatc ggcggatcag cggtcgcccgc gcccagcggc 780
gaaatcctgg caaaagcagc cggcgaggc gatgaagtctg tcgttgtgaa agcagacatc 840
gacatggc a gcccctataa g gaaagcgtc ttcgacttcg cccatcg gcccggc 900
0 gcatacggca tcatcgccga aaggaaaggc cggggcgc cactgcccgt cccgttcaac 960
gtgaatgact aaggatccga aggatata catatggatg caaagctact gttggcggc 1020
actattgttt cctcgaccgg caaaatccga gcccacgtgc tgattaaaaa cggcaaagtc 1080
25 gccgctgtcg gcatgctgga cggcgcacg cggacacag ttgagcgggt tgactgcac 1140
ggcaaatacg tcatgcccgg cggtatcgac gttcacaccc acatcgactc cccctcatg 1200
30 gggaccacca cggccgatga ttttgcagc ggaacgattt cagccgctac cggcggaaaca 1260
acgaccatcg tcgatttcgg acagcagctc gccggcaaga acctgctgga atccgcagac 1320
35 ggcaccacca aaaaggcgc gggaaatcc gtcattgatt acggcttcca tatgtgcgtg 1380
acgaacctct atgacaattt cgattccat atggcagaac tgacacagga cggaaatctcc 1440
agttcaagg tottcatggc taccgcggc agcctgtatc tcaacgcacgg cgaactgttc 1500
45 gacatcctca agggagtcgg ctccagcggt gccaactat gcgtccacgc agagaacggc 1560
gacgtcatcg acaggatcgc cggcgcaccc tacggccaaag gaaaaaccgg gcccgggacc 1620
cacgagatcg cacgcccggc ggaatcgaa gtcgaagcag tcagccggc catcaagatc 1680
50 tcccggatgg ccgaggtgcc gctgtatttc gtgcatttt ccacccagg ggcgtcgag 1740
gaagtagctg cgcgcagat gacaggatgg ccaatcagcg cggaaacgtg cacccactac 1800
55 ctgtcgctga gccgggacat ctacgaccag cgggattcg agccggccaa agctgtcctc 1860
acaccacgc tgcgacaca ggaacaccag gacgcgttgt ggagaggcat taacaccgt 1920
gcgctcagcg tcgtcagttc cgaccactgc cccttctgtt ttgagaaaaa gcagcggatg 1980
ggggcagatg acttccggca gatccccaaac ggcggggcccg gctggagca ccgaatgctc 2040
gtgtatgatg agaccggtgt cgccgaagga aaaatgacga tcgagaaatt cgtcgaggtg 2100

actgccgaga acccgccaa gcaattcgat atgtacccga aaaagggAAC aattgcacccg 2160
ggctccgatg cagacatcat cgtggcgac cccaacggaa caaccctcat cagtgcgcac 2220
5 acccaaaaac aaaacatgga ctacacgctg ttcgaaggct tcaaaatccg ttgctccatc 2280
gaccaggtgt tctcgctgg cgacctgatc agcgtcaaag gcaaatatgt cggcaccgc 2340
ggccgcggcg aattcatcaa gcggagcgct tggagccacc cgcaagtcga aaaataaaag 2400
10 cttggctgtt ttggcgatg agagaagatt ttcagcctga tacagattaa atcagaacgc 2460
agaagcggtc tgataaaaaca gaatttgccct ggccgcagta ggcgggtggt cccacctgac 2520
15 cccatgccga actcagaagt gaaacgcccgt agcgccgatg gtagtgtggg gtctcccat 2580
gcgagagtag ggaactgcca ggcataataaaacgaaag gctcagtcga aagactggc 2640
cttcgtttt atctgtgtt tgtcggtgaa cgctctcctg agtaggacaa atccgcggg 2700
0 agcggatttg aacgttgcga agcaacggcc cggagggtgg cggcaggac gcccgcata 2760
aactgccagg catcaaaatta agcagaaggc catcctgacg gatggcctt ttgcgtttct 2820
25 acaaactctt ttgttttattt ttctaaatac attcaaatat gtatccgctc atgagacaat 2880
aacccctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc 2940
gtgtcgccct tattccctt tttgcggcat tttgccttcc tgttttgct cacccagaaa 3000
30 cgctggtaaaatgat gctgaagatc agttgggtgc acgagtgggt tacatcgaac 3060
tggatctcaa cagcggtaaag atccttgaga gtttcgccc cgaagaacgt tttccaatga 3120
35 tgagcacttt taaaaggcttg ctatgtggcg cggattatc ccgtgttgac gccgggcaag 3180
agcaactcgg tcgcccata cactattctc agaatgactt gttgagttac tcaccagtca 3240
cagaaaaagca tcttacggat ggcatacgac taagagaatt atgcagtgc gccaataacca 3300
tgagtgataa cactgcggcc aacttacttc tgacaacgt cggaggaccg aaggagctaa 3360
ccgctttttt gcacaacatg gggatcatg taactgcct tgatcggttgg gaaccggagc 3420
45 tgaatgaagc cataccaaac gacgagcgtg acaccacgt gcctgttagca atggcaacaa 3480
cgttgcgcaa actattaact ggcataactc ttactctac ttcccgccaa caattaatag 3540
actggatgga ggcggataaa gttgcaggac cacttctgac ctggccctt ccggctggct 3600
50 ggtttattgc tgataaaatct ggagccgggtg agcgtgggtc tcgcgttac attgcagcac 3660
tggggccaga tggtaagccc tcccgatcg tagttatcta cacgacgggg agtcaggcaa 3720
55 ctatggatga acgaaataga cagatcgctg agataggtgc ctcactgatt aagcattgg 3780
aactgtcaga ccaagttac tcataatatac ttttagattga tttaaaaactt catttttaat 3840
ttaaaaaggat ctaggtgaag atcccttttataatctcat gacccaaatc ccttaacgtg 3900

agttttcggtt ccactgagcg tcagaccccg tagaaaagat caaaggatct tcttgagatc 3960

ctttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accacccgcta ccagcggtgg 4020

5 tttgtttgcc ggatcaagag ctaccaactc ttttccgaa ggtaactggc ttcagcagag 4080

cgcagataacc aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact 4140

10 ctgttagcacc gcctacatac ctcgctctgc taatcctgtt accagtggct gctgccagt 4200

gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc 4260

15 ggtcgggctg aacgggggggt tcgtgcacac agcccagctt ggagcgaacg acctacaccg 4320

aactgagata cctacagcgt gagctatgag aaagcgccac gcttccgaa gggagaaaagg 4380

cggacaggtt tccggtaagc ggcagggtcg gaacaggaga ggcacacgagg gagcttccag 4440

40 ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga cttgaggcgtc 4500

gattttgtg atgctcgtca gggggcgga gcctatggaa aaacgccagc aacgcggcct 4560

45 ttttacggtt cctggccttt tgctggcctt ttgctcacat gtttttcct gcgttatccc 4620

ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct cggccagcc 4680

gaacgaccga ggcgcagcgtc agtgtgagcg aggaagcgga agagcgcctg atgcggatt 4740

50 ttctcccttac gcatctgtgc ggtatttcac accgcatata tggtcactc tcagtacaat 4800

ctgctctgtat gcccatagt taagccagta tacactccgc tatcgctacg tgactggcgtc 4860

atggctgcgc cccgacaccc gccaacaccc gctgacgcgc cctgacgggc ttgtctgctc 4920

55 ccggcatccg cttacagaca agctgtgacc gtctccggga gctgcatttg tcagaggttt 4980

tcaccgtcat caccgaaacg cgcgaggcag ctgcggtaaa gctcatcagc gtggcgtga 5040

agcgatttcac agatgtctgc ctgttcatcc gcgtccagct cggtgagttt ctccagaagc 5100

gttaatgtct ggcttctgtat aaagcgggcc atgttaaggg cggttttttc ctgtttggtc 5160

acttgatgcc tccgtgtaag gggaaatttc tgttcatggg gttaatgata ccgatgaaac 5220

55 gagagaggat gtcacgata cgggttactg atgatgaaca tgcccggtt ctggAACgtt 5280

gtgagggtaa acaactggcg gtatggatgc ggcgggacca gagaaaaatc actcagggtc 5340

50 aatgccagcg ctgcgttaat acagatgtag gtgttccaca gggtagccag cagcatctg 5400

cgatgcagat ccggAACata atggcgcagg gcgcgtactt ccgcgtttcc agactttacg 5460

aaacacggaa accgaagacc attcatgtt tgctcaggt cgcagacgtt ttgcagcagc 5520

agtcgcgttca cggtcgctcg cgtatcggtt attcattctg ctaaccagta aggcaacccc 5580

55 gccagcctag ccgggtccctc aacgacagga gcacgatcat gcgcacccgt ggccaggacc 5640

caacgctgcc cgagatgcgc cgctgcggc tgctggagat ggcggacgcg atggatatgt 5700
tctgccaagg gttggttgc gcattcacag ttctccgcaa gaattgattg gctccaattc 5760
5 ttggagtggt gaatccgtta gcgaggtgcc gccggcttcc attcaggtcg aggtggcccg 5820
gctccatgca cgcgcacgca acgcggggag gcagacaagg tataggcgg cgcctacaat 5880
ccatgccaac ccgttccatg tgctcgccga ggcggcataa atcgccgtga cgatcagcgg 5940
10 tccagtgatc gaagtttaggc tggtaagagc cgcgagcgat ccttgaagct gtccctgatg 6000
gtcgtcatct acctgcctgg acagcatggc ctgcaacgcg ggcattccga tgccgccgga 6060
15 agcgagaaga atcataatgg ggaaggccat ccagcctcgc gtcgcgaacg ccagcaagac 6120
gtagcccagc gcgtcgcccg ccatgcccgc gataatggcc tgcttctcgc cgaaacgttt 6180
ggtggccggga ccagtgacga aggcttgagc gagggcgtgc aagattccga ataccgcaag 6240
cgacaggccg atcatcgatc cgctccagcg aaagcggtcc tcgcccaaaa tgacccagag 6300
cgctgcccgc acctgtccatc cgagttgcgt gataaagaag acagtcataa gtgcggcgac 6360
25 gatagtcatg ccccgccccc accggaagga gctgactggg ttgaaggctc tcaaggcat 6420
cggtcgacgc tctcccttat gcgactcctg cattaggaag cagcccgatc gtaggttgag 6480
gcccgttgc accgcccggc caaggaatgg tgcatgctcg atggctacga gggcagacag 6540
30 taagtggatt taccataatc ccttaattgt acgcaccgct aaaacgcgtt cagcgcgatc 6600
acggcagcag acaggtaaaa atggcaacaa accaccctaa aaactgcgcg atcgccctg 6660
35 ataaaatttta accgtatgaa tacctatgca accagagggc acaggccaca ttaccccccac 6720
ttaatccact gaagctgcca ttttcatgg tttcaccatc ccagcgaagg gccatgcgt 6780
catcgaaatt aatacgacga aattaatacg actcaactata gggcaattgc gatcaccaca 6840
attcagcaaa ttgtgaacat catcacgttc atcttccct gttgccaat ggcccatattt 6900
cctgtcagta acgagaaggt cgcgaaattca ggcgctttt agactggtcg
taatgaac 6958